DTIC® has determined on 05/24/2010 that this Technical Document has the Distribution Statement checked below. The current distribution for this document can be found in the DTIC® Technical Report Database.

☒ **DISTRIBUTION STATEMENT A.** Approved for public release; distribution is unlimited.

☐ **© COPYRIGHTED**; U.S. Government or Federal Rights License. All other rights and uses except those permitted by copyright law are reserved by the copyright owner.

☐ **DISTRIBUTION STATEMENT B.** Distribution authorized to U.S. Government agencies only (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office)

☐ **DISTRIBUTION STATEMENT C.** Distribution authorized to U.S. Government Agencies and their contractors (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office)

☐ **DISTRIBUTION STATEMENT D.** Distribution authorized to the Department of Defense and U.S. DoD contractors only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

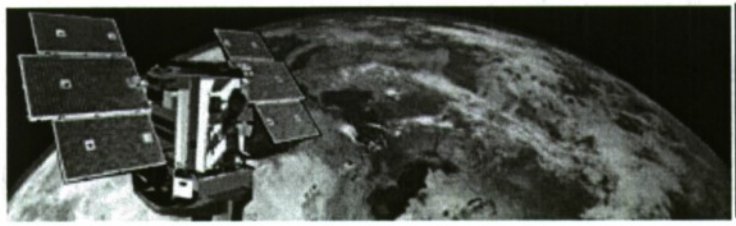
☐ **DISTRIBUTION STATEMENT E.** Distribution authorized to DoD Components only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

☐ **DISTRIBUTION STATEMENT F.** Further dissemination only as directed by (inserting controlling DoD office) (date of determination) or higher DoD authority.

Distribution Statement F is also used when a document does not contain a distribution statement and no distribution statement can be determined.

☐ **DISTRIBUTION STATEMENT X.** Distribution authorized to U.S. Government Agencies and private individuals or enterprises eligible to obtain export-controlled technical data in accordance with DoDD 5230.25; (date of determination). DoD Controlling Office is (insert controlling DoD office).

NASA TECHNICAL REPORTS SERVER (NTRS)



Search NTRS

NAVIGATION SEARCH OPTIONS
Collection > NASA

NASA Center > Johnson Space Center

Publication Year > 1991-2000 > 1993

Subject > M-O > Man/system Technology And Life Support

Availability Options > Online > PDF

Item/Media Type > Collected Works

Item/Media Type > NASA Report > Technical Memorandum (TM)

Note: Start a new navigation search by selecting a link above

SEARCH NTRS

+ Ba
+ Pri

Title:	**Crew interface analysis**: Selected articles on space human research, 1987 - 1991
Author(s):	NONE
Abstract:	As part of the Flight Crew Support Division at NASA, the Cre Analysis Section is dedicated to the study of human factors ir manned space program. It assumes a specialized role that fo answering operational questions pertaining to NASA's Space and Space Station Freedom Programs. One of the section's l contributions is to provide knowledge and information about l capabilities and limitations that promote optimal spacecraft a design and use to enhance crew safety and productivity. The provides human factors engineering for the ongoing missions proposed missions that aim to put human settlements on the Mars. Research providing solutions to operational issues is tl objective of the Crew Interface Analysis Section. The studies such subdisciplines as ergonomics, space habitability, man-c interaction, and remote operator interaction.
NASA Center:	Johnson Space Center
Publication Date:	Jul 1, 1993
Document Source:	CASI
Online Source:	View PDF File
Document ID:	19940019712
Accession ID:	94N24185
Publication Information:	Number of Pages = 137
Report Number:	NAS 1.15:104756; NASA-TM-104756; S-697
Price Code:	A07
Related Information:	Sponsored by NASA. Washington
Keywords:	ASTRONAUTS; BIOASTRONAUTICS; HABITABILITY; HUM FACTORS ENGINEERING; HUMAN PERFORMANCE; HUM COMPUTER INTERFACE; MANNED SPACE FLIGHT; REM CONTROL; SPACE HABITATS; SPACECRAFT DESIGN; SPACECREWS; TELEOPERATORS; COGNITIVE PSYCHO LUNAR BASES; MENTAL PERFORMANCE; SAFETY FACT SPACE FLIGHT STRESS; SPACE PROGRAMS; SPACE SI SPACE STATION FREEDOM;
Accessibility:	Unclassified; No Copyright; Unlimited; Publicly available;

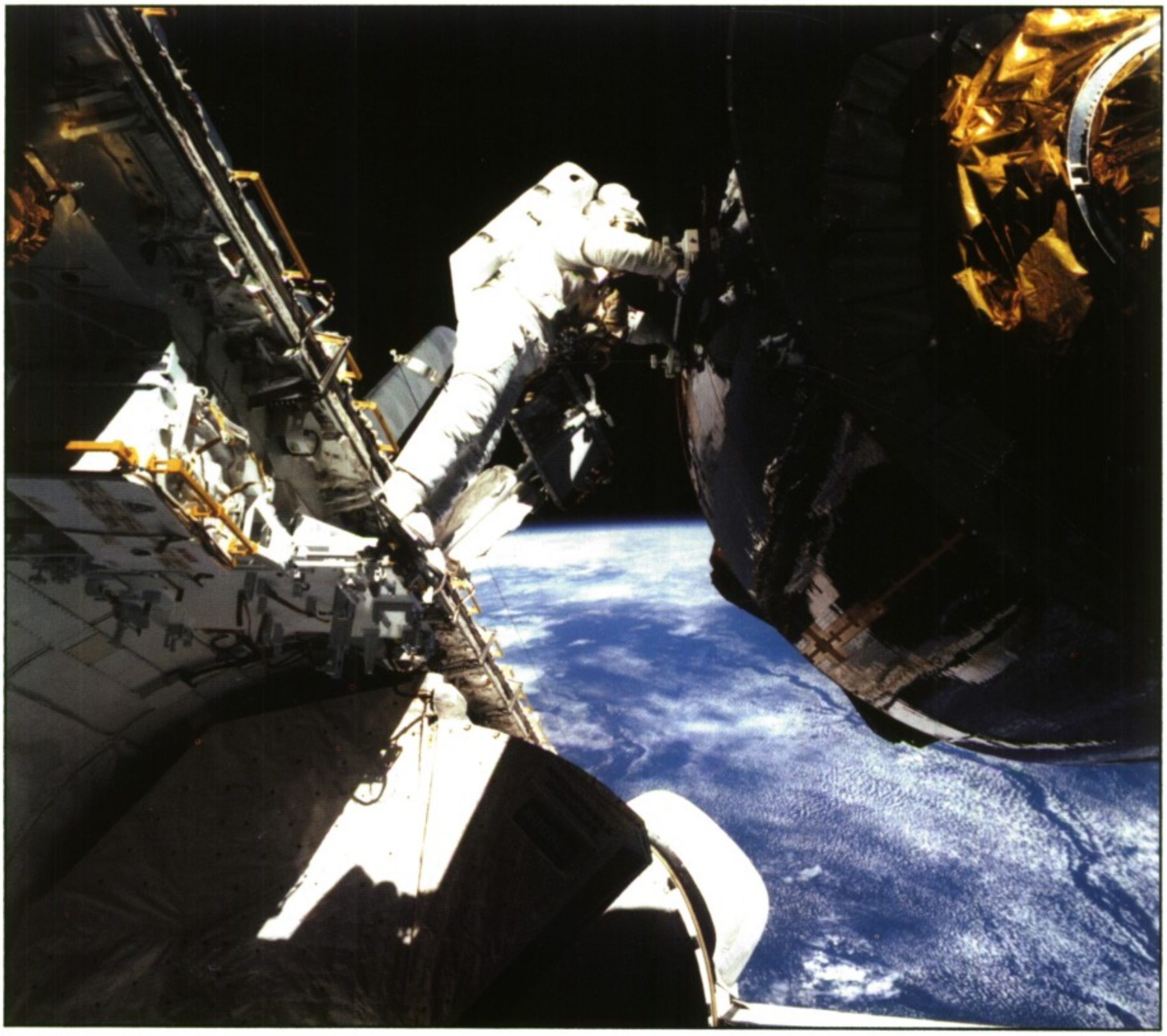
Updated/Added to NTRS: 2006-08-02

+

TR-01-644

NASA Technical Memorandum 104756

Crew Interface Analysis: Selected Articles on Space Human Factors Research, 1987 - 1991



July 1993

NASA

NASA Technical Memorandum 104756

Crew Interface Analysis: Selected Articles on Space Human Factors Research, 1987 - 1991

National Aeronautics and Space Administration
Lyndon B. Johnson Space Center
Houston, Texas

July 1993

PREFACE

This document contains human factors articles written between 1987 and 1991 and an acknowledgement with publication information. From its inception, this document was intended to focus on presenting a cross section of the Johnson Space Center's Crew Interface Analysis Section's work in progress.

These articles were generated in the course of everyday work and were not written specifically for this document. They provide a sampling of the work in progress in the Anthropometry and Biomechanics Laboratory, the Graphics Analysis Facility, the Human-Computer Interaction Laboratory, the Remote Operator Interaction Laboratory, the Lighting Environment Test Facility, and the Task Analysis Laboratory. The articles are organized by topic area and not laboratory thus emphasizing the interdisciplinary and integrated approach to research adopted by the section.

It is hoped that these articles demonstrate the versatility and professionalism of the section, and, in so doing, instill a broad appreciation for the importance of human factors engineering in the design of human-operated systems with particular emphasis on space systems and missions.

CONTENTS

SECTION THREE - *Remote Operator Interaction*

SECTION FOUR - *Ergonomics*

ACKNOWLEDGEMENTS

The following material reprinted with permission by Elsevier Science Publishers B.V.

Chandlee, G. O., Smith, R. L., and Wheelwright, C. D. Illumination requirements for operating a space remote manipulator. *Ergonomics of Hybrid Automated Systems*, volume 1, pp. 241-248.

The following reprinted with permission © 1990/1988 Society of Automotive Engineers, Inc.

Bagian, J. P., Greenisen, M. C., Schafer, L. E., Probe, J. D., and Krutz, R. W. (1990) Reach Performance While Wearing the Space Shuttle Launch and Entry Suit During Exposure to Launch Accelerations. SAE Technical Paper Series No. 901357.

Marmolejo, J. A. and Shepherd, C. K. (1988) A simulation system for Space Station extravehicular activity. SAE Technical Paper Series No. 881104

The following reprinted with permission. Copyright 1990, Association for Computing Machinery, Inc.

Gillan, D. J. and Breedin, S. D. (1990) Designers' models of the human-computer interface. *Proceedings of the Association for Computing Machinery's Computer Human Interaction (CHI) Conference*. New York, NY:ACM.

The following reprinted with permission by the International Astronautical Federation.

Goldsberry, B. S. Lippert, B. O., McKee, S. D., Lewis, J. L, Jr., and Mount, F. E. (1989) Using computer graphics to design Space Station Freedom viewing. *Proceedings of the International Astronautical Federation (IAF) Congress of Malaga,* Vol. 22 of Acta Astronautica.

The following reprinted with permission from Proceedings of the Human Factors Society Annual Meeting. Copyright (1987/1988/1989/1990) by The Human Factors Society, Inc. All rights reserved.

Bierschwale, J. M., Sampaio, C. E., Stuart, M. A., and Smith, R. L. (1989) Speech versus manual control of camera functions during a telerobotic task, *Proceedings of the Human Factors Society 33rd Annual Meeting*, pp.134-138.

Chandlee, G.O. and Woolford, B. (1988). Previous experience in manned space flight: a survey of human factors lessons learned. *Proceedings of the Human Factors Society 32nd Annual Meeting*. pp. 49-52.

Micocci, A. J. (1987) Use of infrared telemetry as part of a nonintrusive inflight data collection. *Proceedings of the Human Factors Society 31st Annual Meeting*, pp. 796-799.

Shepherd, C. K. (1990) Performing speech recognition research with Hypercard. *Proceedings of the Human Factors Society 34th Annual Meeting*, pp. 415-418.

Smith, R. L. and Stuart, M. A. (1989) The effects of spatially displaced visual feedback on remote manipulator performance. *Proceedings of the Human Factors Society 32nd Annual Meeting*. pp. 1430-1434.

Stuart, M. A., Smith, R. L., and Moore, E. P. (1988) Guidelines for the use of programmable display pushbuttons on the Space Station's telerobot control panel. *Proceedings of the Human Factors 32nd Annual Meeting*, pp. 44-48.

The following articles were printed as NASA Conference Publications.

Gillan, D. J., Lewis, R., and Rudisill, M. (1989) Process and representation in graphical displays. *Proceedings of SOAR (Space Operations, Automation, and Robotics), the Third Annual Workshop on Automation and Robotics*, pp. 661-665.

Gugerty, L. (1990) The use of analytical models in human-computer interface design. *Proceedings of SOAR (Space Operations, Automation, and Robotics), the Third Annual Workshop on Automation and Robotics*, pp. 574-581.

O'Neal, M. R. and Manahan, M. K. (1990) Spacecraft crew procedures from paper to computers. *Proceedings of SOAR (Space Operations, Automation, and Robotics), the Third Annual Workshop on Automation and Robotics*, pp. 595-600.

Probe, J. D. (1989) Quantitative assessment of human motion using video motion analysis. *Proceedings of SOAR (Space Operations, Automation, and Robotics), the Third Annual Workshop on Automation and Robotics*, pp. 155-157.

Sampaio, C. E., Bierschwale, J. D., Fleming, T. F., and Stuart, M. A. (1990) Space Station Freedom coupling tasks: an evaluation of their telerobotic and EVA compatibility. *Proceedings of SOAR (Space Operations, Automation, and Robotics), the Third Annual Workshop on Automation and Robotics*, pp. 179-183.

Stuart, M. A. and Smith, R. A. (1988) Simulation of the human telerobotic interface on the Space Station. *Proceedings of SOAR (Space Operations, Automation, and Robotics), the Third Annual Workshop on Automation and Robotics*, pp. 321-326.

Woolford, B. Pandya, A. K. and Maida, J. (1990) Development of biomechanical models for human factors evaluations. *Proceedings of SOAR (Space Operations, Automation, and Robotics), the Third Annual Workshop on Automation and Robotics*, pp. 552-556.

INTRODUCTION

This document is a collection of studies that illustrate recent work conducted by members of the Crew Interface Analysis Section (CIAS). It represents an advancement of knowledge concerning integrating humans into the spaceflight environment. The studies include such subdisciplines as ergonomics, space habitability, human-computer interaction, and remote operator interaction. The CIAS is dedicated to human factors study in the manned space program.

The CIAS is one of the Flight Crew Support Division's seven sections in NASA Johnson Space Center's Space and Life Sciences Directorate. The Flight Crew Support Division is concerned with human factors issues in spaceflight and the CIAS assumes a more specialized role that focuses on research, evaluation, development, and preliminary design for advanced projects. A key contribution of the CIAS is to provide knowledge and information about the capabilities and limitations of the human so that spacecraft systems and habitats are optimally designed and used, and crew safety and productivity are enhanced.

The section provides human factors engineering for the Space Shuttle program, the Space Station Freedom program, and advanced programs that have the goal of human Lunar and Martian exploration. The CIAS also contributes to the Space and Life Sciences Directorate's goal to serve as a focal point of excellence for the development and implementation of procedures, hardware, and science payloads that relate directly to the health, safety, and performance of humans in space.

A goal of the CIAS is to increase program commitment to designing for efficient human productivity in the space environment. Ongoing research and development activities enhance the understanding of crew capabilities and ensure the best use of humans in the manned space programs. The application of human factors principles to spacecraft and mission design will result in optimally engineered systems, and contribute to the achievement of NASA's goals. The following set of articles illustrates how these principles have already benefited the space program.

Human-Computer Interface

Principles of human-computer interaction are researched and developed to be applied to the design of computer interfaces on NASA space missions. Interfaces between the human operator and the computer system are evaluated to measure and record parameters such as formats for displays, input/output devices and workstation layouts.

SPACECRAFT CREW PROCEDURES FROM PAPER TO COMPUTERS

Michael O'Neal and Meera Manahan
Lockheed Engineering and Sciences Company

Research directed by Marianne Rudisill, Manager, Human Computer Interaction Lab, NASA JSC.

INTRODUCTION

Large volumes of paper are launched with each Space Shuttle mission that contain step-by-step instructions for various activities that are to be performed by the crew during the mission. These instructions include normal operational procedures and malfunction or contingency procedures and are collectively known as the Flight Data File, or FDF. An example of nominal procedures would be those used in the deployment of a satellite from the Space Shuttle; a malfunction procedure would describe actions to be taken if a specific problem developed during the deployment.

A new Flight Data File and associated system is being created for Space Station Freedom. The system will be called the Space Station Flight Data File, or SFDF. NASA has determined that the SFDF will be computer-based rather than paper-based for reasons including the following:

- The long duration of the Space Station program precludes one-time launch of all crew procedures.
- Repeated launch of crew procedure segments is not cost effective since each pound of launch weight costs approximately $20,000.
- Large amounts of manual effort are required to create, edit, and maintain paper-based crew procedures.
- Changes made after crew procedure printing require annotation of each individual copy, a time-consuming and error-prone process.
- The time involved in implementing and delivering approved Space Station crew procedure changes or updates in a paper-based system would be significant, including scheduling of resources on a Space Shuttle flight.

The main components of interest in a Human-Computer Interface (HCI) include the information available on the screen at any given time, how to change the quantity or content of the information present on the screen, how the information is organized, and how the user interacts with the displayed information. Designing an effective HCI is an important step in developing a viable computer-based crew procedure system for reasons including the following:

- An effective HCI will allow faster, more accurate crew interaction with spacecraft computer procedure systems.
- The HCI will facilitate the crew's monitoring of other spacecraft computer systems while performing crew procedures.
- The HCI will allow the crew to easily verify procedure steps performed by the computer system as procedure automation increases.
- A context- and user- sensitive help and annotation system within the HCI will allow the user to rapidly and efficiently access this type of information while performing the procedures.
- The effective HCI will provide rapid, easy access to required supporting information such as procedure reference items.
- The development of a standard HCI across all crew procedures will lessen the amount

of cross-training required for different types of procedures and will thus lessen the amount of errors made during procedures.

The research project described in this paper uses human factors and computer systems knowledge to explore and help guide the design and creation of an effective HCI for computer-based spacecraft crew procedure systems. The research project includes the development of computer-based procedure system HCI prototypes and a test-bed including a complete system for procedure authoring, editing, training, and execution to be used for experiments that measure the effectiveness of HCI alternatives in order to make design recommendations.

CREW PROCEDURE TASKS AND USERS

Many different tasks are required to create and maintain a spacecraft crew procedure system. Procedures must be created by personnel familiar with the tasks in question and by procedure authors and editors. The crew responsible for performing the procedures must be trained in how to use the procedures. Training crew personnel to be familiar with off-nominal procedures is also required so the procedures can be used quickly and effectively if needed during a mission. The main procedure task will be its actual performance during a mission, including assistance and adaptation to changing conditions if necessary. If a procedure is used repeatedly during one or more missions, changes to the procedure may be required to correct inefficiencies or errors, and current versions of such procedures must be maintained and distributed to all appropriate personnel.

Personnel groups responsible for specific crew procedure tasks represent different user groups of the crew procedure system. Procedure authors create the procedures, assuring that they correctly describe the work to be performed and that they conform to a standard procedural format (e.g., FDF or SFDF); they are also involved in scheduling procedures during a mission to create mission plans and crew member short-term plans. Authors may also work with individual payload specialists or experimental scientists. Trainers review the procedures with the crew members who will perform the tasks; comments or problems with procedure details or clarity are reported to procedure authors or editors for correction. Crew members are involved with actual procedure performance, training, and correction or editing if required. Mission control personnel assist in scheduling procedures, working with the crew during the mission, and in monitoring the mission plan and short-term plans. Experimental investigators and payload specialists are involved in creation and execution of those procedures relevant to their experiment or payload. Procedure editors are also responsible for updating and distributing required procedure changes found during training or execution.

An effective computer-based crew procedure system, and an effective HCI to this system, must take into account the full range of tasks and users of the procedure system. In particular, a common interface that can be created by authors and used by trainers, crew members, and mission control personnel will contribute to faster, more accurate interaction with crew procedures.

PROJECT GOALS

The final goal of the current research is to create HCI design guidelines that can be used for spacecraft crew procedures and other computer systems that display procedural information to procedure users. These guidelines should lead to faster, more accurate user interaction with procedural information on a computer.

The first step in the project is a review of available literature on computer presentation of procedural material and the evaluation of the current paper-based FDF procedure system for Space Shuttle. With this information, key issues are identified and their role in the research outlined. Using background information and human factors and computer system knowledge, alternative interfaces are

created via prototypes. These prototypes are then evaluated by the various users of crew procedures listed above. Experiments are then performed using different presentation and interaction techniques; these experiments provide specific data on the relative speed and accuracy of procedure tasks using different interfaces. Comments from prototypes and results and conclusions from interface experiments are then compiled into human-computer interface guidelines for presentation and interaction with spacecraft crew procedures.

CREW PROCEDURE ISSUES

There are both advantages and disadvantages of moving from a paper-based to a computer-based crew procedure system. The current research project addresses these issues as they relate to the human-computer interface of the system. Advantages of using a computer will be utilized while disadvantages will be addressed and minimized.

COMPUTER ADVANTAGES

Having a computer system behind the interface to a crew procedure system offers many advantages. By monitoring related onboard systems, the computer system can automatically perform many procedure steps that require simple status verification (e.g., "Check that switch F6 is ON"), thus reducing the time required to perform the procedure. A training mode is now feasible so that the crew member can practice using the procedure in exactly its final form with the exception that system actions are not actually performed; training and execution modes for the same procedure will increase the effectiveness of training. Personal annotation files can be attached to each procedure, thus allowing each crew member to create and refer to individual notes during both training and execution of procedures; these notes will be available whenever and wherever the crew member uses the procedure. The computer-based procedure system can coordinate with other spacecraft computer systems, providing easier transitions to and from other systems. The computer-based help system can adapt to both the user of the procedure and the context in which the procedure is being performed. The amount of detail (i.e., the prompt level) of the procedure can change for different users and situations. Finally, expert systems can be integrated into the procedure system, thus providing a more intelligent interface to crew procedures.

COMPUTER DISADVANTAGES

When procedural information is presented on a computer screen, the context of the information presented typically seems more limited than with a page of paper, although the actual amount of information present on a computer screen may or may not be smaller. There is less context information on where the current screen of information fits into the overall system; in a book, the location of the page in the overall book is an example of available context data. This issue will be addressed in the HCI to the computer-based system by generating and evaluating ideas to provide additional context information (e.g., screen number, screen position in overall outline, etc.).

In a complex computer system such as the onboard Data Management System (DMS) for Space Station *Freedom*, many levels of subsystems are present. The inability to rapidly navigate among the systems and subsystems can be a serious detriment to overall performance. This issue will be addressed in the HCI to the computer-based system by generating and evaluating ideas to provide information on current position within the system hierarchy and to provide tools to rapidly and directly move between subsystems either during or after a computer task.

RESEARCH FOUNDATIONS

Initially, a review of NASA literature on computer presentation of procedural information was completed. Information on work performed at MITRE for the Procedure Formatting System (PFS) project was received and prototypes were viewed (Johns 1987 and 1988, Kelly 1988). Previous research in the Human-Computer Interaction Laboratory (HCIL) of the NASA Johnson Space Center was reviewed, and results from experiments

on procedure context and format will be incorporated into the current research project (Desaulniers, Gillan, and Rudisill 1988 and 1989). Coordination is in progress with the Mission Operations Directorate (MOD) at the NASA Johnson Space Center, as described below.

PROJECT STATUS

CURRENT PROJECT PROTOTYPES

Prototype development is in progress for two Space Shuttle experiments. The procedures were selected for prototyping due to their similarity to typical research that will be conducted on Space Station *Freedom* since Space Station procedures are not yet available. The two prototypes will also use different HCI approaches.

The first system is a computer-based prototype of a middeck experiment, Polymer Morphology (PM), that was performed on Space Shuttle mission STS-34. The PM experiment consists of four procedures (set up, sequence initiation, sample check, and stowage) and six procedure reference items (interconnection overview, keystroke definitions, window definitions, notebook, sequences, and worksheets). The prototype is created within the framework of the Space Station basic screen layout being developed by the DMS development team. Included in this prototype is an initial version of an Interface Navigation Tool developed at the HCIL that is currently being reviewed by the DMS team. Initial versions of the six reference items have been created. Development of the interface for the four procedures of the experiment is in progress.

The second system is a computer-based prototype of an expert system for medical experiments to be performed on two upcoming Space Shuttle missions. The system, Principal Investigator in a Box, or [PI], will include an expert system. The motivation for this medical expert system is to provide the capability to perform medical experiments with minimum ground control or support. A separate HCIL research project is in progress to study the interface as it relates to the expert system, and this research will be coordinated with the current research which examines the same interface from the viewpoint of presentation of the procedures. The [PI] interface is being modified for the Space Station basic screen layout and will be evaluated as an alternative HCI design for crew procedures.

CURRENT PROJECT EXPERIMENTS

As discussed above, the current procedures research will include the performance of experiments to gather specific data to support HCI guidelines for computer presentation of procedures. These experiments will begin as specific questions arise from the creation and analysis of HCI prototypes. The experiments will use subjective comments and speed and accuracy measurements to provide data for comparing different HCI alternatives. The experimental test-bed will include a complete system for procedure authoring, editing, training, and execution that will allow HCI alternatives to be easily generated and compared.

COOPERATIVE WORK

In addition to continuing work with the MITRE PFS system, two cooperative projects with the NASA Johnson Space Center Mission Operations Directorate (MOD) are in the planning stages. Research will be performed in the HCIL to assist MOD in creating procedure standards for SFDF. Studies and experiments will be performed to provide human factors input into the standards created. Also, procedure authoring and execution software being developed within MOD will be evaluated from a human factors and HCI perspective.

FUTURE RESEARCH ISSUES

The current research project will continue to explore human factors issues relevant to the interface to electronic spacecraft crew procedures. The effect on the cognitive workload of the procedure users will be examined, with the goal of reducing this workload through automation. The allocation

of procedure tasks between the user and the computer system will also be examined. Creating an interface that is adaptable to changing environments will be explored, including the method and user aids available during interruption and resumption of procedures. Research will also be performed on the use of the same computer interface during both training and execution of procedures.

CONCLUSION

Spacecraft crew procedures are increasingly being computerized, as in NASA's Space Station *Freedom* program. The human interface to these computer-based crew procedure systems is an important component, and research into improving the interface will provide faster and more accurate human interaction with the computer. The current research project uses prototypes and experiments to explore and help guide the design and creation of the human-computer interface for spacecraft crew procedure systems such as the Space Station. Prototype and experiment development is currently in progress. Issues relevant to human interaction with procedures will continue to be researched within the HCIL and in cooperation with other crew procedures researchers and developers.

ACKNOWLEDGEMENTS

This research was funded by the National Aeronautics and Space Administration, Office of Aeronautics and Exploration Technology, through contract NAS9-17900 to Lockheed Engineering and Sciences Company. The research was performed at the Johnson Space Center Human-Computer Interaction Laboratory.

REFERENCES

1. *Desaulniers, D. R., Gillan, D. J., and Rudisill, M. R. (1988).* The Effects of Format in Computer-Based Procedure Displays. *NASA Technical Report. Houston, Texas: Lockheed Engineering and Sciences Company.*

2. *Desaulniers, D. R., Gillan, D. J., and Rudisill, M. R. (1989).* The Effects of Context in Computer-Based Procedure Displays. *NASA Technical Report. Houston, Texas: Lockheed Engineering and Sciences Company.*

3. *Desaulniers, D. R., Gillan, D. J., and Rudisill, M. R. (1988).* A Comparison of Flowchart and Prose Formats for the Presentation of Computer-Based Procedural Information. *NASA Technical Report. Houston, Texas: Lockheed Engineering and Sciences Company.*

4. *Johns, G. J. (1987).* Flight Data File for the Space Station, Volume I, Baseline Definition, and Volume II, Baseline Functional Requirements. *(MITRE Report No. MTR10019). Houston, Texas: The MITRE Corporation.*

5. *Johns, G. J. (1988).* Dynamic Display of Crew Procedures for Space Station, Volume I, Baseline Definition, and Volume II, Functional Requirements. *(MITRE Report No. MTR-88D033). McLean, Virginia: The MITRE Corporation.*

6. *Kelly, C. M. (1988).* Conceptual Definition for a Flight Data File Automated Control and Tracking System. *(MITRE Report No. MTR-88D0017). McLean, Virginia: The MITRE Corporation.*

PROCESS AND REPRESENTATION IN GRAPHICAL DISPLAYS

Douglas J. Gillan
Lockheed Engineering and Sciences Company

Robert Lewis
Rice University

Marianne Rudisill
NASA Lyndon B. Johnson Space Center

INTRODUCTION

To survive and succeed in the world, people have to comprehend both diverse natural sources of information, such as landscapes, weather conditions, and animal sounds, and human-created information artifacts such as pictorial representations (i.e., graphics) and text. Researchers have developed theories and models that describe how people comprehend text (for example, see [8]), but have largely ignored graphics. However, an increasing amount of information is provided to people by means of graphics, as can be seen in any newspaper or news magazine, on television programs, in scientific journals and, especially, on computer displays.

Our initial model of graphic comprehension has focused on statistical graphs for three reasons: (1) recent work by statisticians which provides guidelines for producing statistical graphs (Bertin [2], Cleveland and McGill [4,5] and Tufte [10]) could be translated into preliminary versions of comprehension models, (2) statistical graphs play an important role in two key areas of the human-computer interface — direct manipulation interfaces (see [7] for a review) and task-specific tools for presenting information, e.g., statistical graphics packages, and (3) computer-displayed graphs will be crucial for a variety of tasks for the Space Station *Freedom* and future advanced spacecraft. Like other models of human-computer interaction (see [3] for example), models of graphical comprehension can be used by human-computer interface designers and developers to create interfaces that present information in an efficient and usable manner.

Our investigation of graph comprehension addresses two primary questions — how do people represent the information contained in a data graph and how do they process information from the graph? The topics of focus for graphic representation concern the features into which people decompose a graph and the representation of the graph in memory. The issue of processing can be further analyzed as two questions, what overall processing strategies do people use and what are the specific processing skills required?

GRAPHIC REPRESENTATION

FEATURES OF GRAPHIC DISPLAYS

Both Bertin [2] and Tufte [10] address the features underlying the perception and use of graphs. Bertin [2] focuses on three constructs, (1) "implantation," i.e., the variation in the spatial dimensions of the graphic plane as a point, line, or area; (2) "elevation," i.e., variation in the spatial dimensions of the graphical element's qualities — size, value, texture, color orientation, or shape; and (3) "imposition," i.e. how information is represented, as in a statistical graph, a network, a geographic map, or a symbol. Tufte [10] proposes two features as important for graphic construction, data ink and data density. Tufte describes data ink as "the nonerasable core of a graphic" [10, p. 93] and provides a measure, the data-ink ratio, which is the "proportion of a graphic's ink devoted to the nonredundant display of data information" [10, p. 93]. Data density is the ratio of the number of data points and the areas of the graphic. Tufte's guidelines call for maximizing both the data-ink ratio and, within reason,

the data density, in other words, displaying graphics with as much information and as little ink as possible.

Both Bertin's and Tufte's ideas about the features of data graphs were derived from their experience as statisticians, rather than from experimental evidence. We decided to fill the empirical void concerning the features underlying graphic comprehension. In our first experiment, people simply judged the similarity in appearance and information displayed by all possible pairs of 17 different types of graphs (that is, 136 pairs of graphs). The graphs ranged from the familiar (line graphs, bar graphs, and scatter plots) to the more unusual (star graphs, ray graphs, and stick man graphs). The similarity judgments were analyzed with multivariate statistical techniques, including (1) cluster analysis, which shows the groupings or categories (clusters) that underlie people's judgments about a set of objects and (2) multidimensional scaling (MDS), which shows the linear dimensions underlying people's similarity judgments. The logic of these analyses was that people would cluster graphs and place graphs along dimensions based on the features of the graph [9].

The cluster analyses indicated that people group graphs, at least in part, according to the physical elements of the graphs. Key clusters include graphs in which points were the dominant element (the two types of scatter plots, the range and density graphs), graphs consisting of straight lines (the surface, textured surface, and stacked bar graph), and those consisting of solid areas (the column and bar graphs). The categorization of the graphs according to physical elements agrees generally with Bertin's [2] construct of implantation.

The MDS analyses of the similarity judgments were combined with a factor analysis which resulted in three factors, each consisting of one informational dimension and one perceptual dimension, which accounted for 97% of the data. One factor differentiated perceptually simple graphs (e.g., the bar and line graphs) from perceptually complex graphs (the scatter plots, the 3-dimensional graph, and the surface graphs). A second factor separated graphs for which axes were unnecessary to read the graph (the pie, star, 3-dimensional, and stick man graphs) from those for which the axis contained information (especially the modified scatter plots — the range and density graphs [10]). Finally, the third factor tended to have informationally complex graphs (those with the most data) at one end and informationally simple graphs (those with the least data) at the other end. Accordingly, we hypothesize that people decompose a graph according to its perceptual complexity, figure-to-axes relation, and informational complexity. A subsequent experiment has shown that each of these factors relates to peoples' speed and accuracy in answering questions using these graphs [6].

REPRESENTATION IN MEMORY

The previous section of this paper addressed the features present when a user looks at a graphic. This section addresses the features that the user walks away with. Accordingly, the experiments looked at how a user represents the information from a graphic in memory.

Our research on memorial representation of graphics involved a simple experimental design: Our subjects worked with a set of graphs on one day, then we assessed what they retained about the graphic on a second day. The initial training day consisted of one trial with each of six different graphs during a 30 second trial. For three graphs, the subjects answered questions about the graphs, (e.g., What is the mean of the variables in the graph? and Which has the greater value, variable A or variable B?). For the other three graphs, they identified and drew the perceptual components of the graph, each component in a separate box. For example, in a line graph a subject might draw the points representing each variable, the lines connecting the points, the axes, verbal labels, and numerical labels.

Twenty-four hours after training, we tested the subjects using two different methods. We gave one group of 16 subjects a recognition test in which they looked at 24 different graphs

and had to say whether they had seen precisely that graph during the training session. We constructed the 24 test graphs systematically. Each of the six graphs from the training session were presented during the test. Each training graph had three "offspring" that served as the distractors (or incorrect test stimuli) during the test. One type of distractor contained the same data as the training stimulus, but used a different graph type to display the data (New Graph-Same Data); a second distractor displayed the data using the same type of graph (Same Graph-New Data); the third distractor differed from the training graph in both graph type and data (New Graph-New Data). Perfect recognition would have resulted in 100% yes answers to the training graphs and 0% yes answers to the distractors. A second group of 14 subjects received a recall test in which they were asked to draw the graphs from Day 1 in as much detail as they could remember.

The results showed that people's recognition of the training graphs was very good. They correctly recognized the training graph 88% of the time, with little difference between the graphs used during training in the perceptual task (85% recognition) and those used in the informational task (90% recognition). Although false recognitions of the distractors were low overall (10% yes answers to distractors), the distribution of false recognitions was interesting. Of the 39 false recognitions by the 16 subjects, 29 (74%) were made to the Same Graph-New Data distractor. Friedman test chi-square (2 df) = 10.1, $p < .05$. The high false recognition rate when the same graph type was used (30% false recognitions to that distractor) suggest that the perceptual type of the graph has a strong representation in memory. We found that both training with an informational task and training with a perceptual task yielded similar high proportions of the total false recognitions for the Same Graph-New Data distractor, 77% and 70%, respectively.

The results from the recall test provide even greater support for the hypothesis that the representation of the graph type and certain perceptual features was exceptionally strong. Subjects had good recall for the graph type (71% of the graphs), the presence or absence of axes (71% correct recall of axes), and the perceptual elements (lines, areas, and points) in the graphs (53% correct recall of graph elements). In contrast, recall of information from the graphs was generally poor. For example, subjects had low recall rates for the number of data points in the graph (29% correct recall), the quantitative labels on the axes (10% of the labels), and the verbal labels of the axes and data points (12% of verbal labels). They recalled the correct spatial relations between data points only 22% of the time. In addition to showing the strength of the perceptual representation, these data suggest that the perceptual and informational representations of a graph are independent.

STRATEGIES FOR PROCESSING INFORMATION

Based on formal thinking-aloud protocols, as well as informal discussions with users, we have hypothesized that people use two different types of strategies when processing information from a data graph - an arithmetic, look-up strategy and a perceptual, spatial strategy. With the arithmetic strategy, a user treats a graph in much the same way as a table, using the graph to locate variables and look up their values, then performing the required arithmetic manipulations on those variables. In contrast, the perceptual strategy makes use of the unique spatial characteristics of the graph, comparing the relative location of data points.

We have hypothesized that users apply the strategies as a function of the task. Certain tasks appear to lend themselves better to one strategy than another. Answering a comparison question like "Which is greater, variable A or B?" would probably be answered rapidly and with high accuracy by comparing the spatial location of A and B. In contrast, a user answering the question "What is the difference between variables A and B?" about a line graph might be able to apply the perceptual strategy, but would be able to determine the answer more easily and accurately with the arithmetic strategy. In addition, we propose

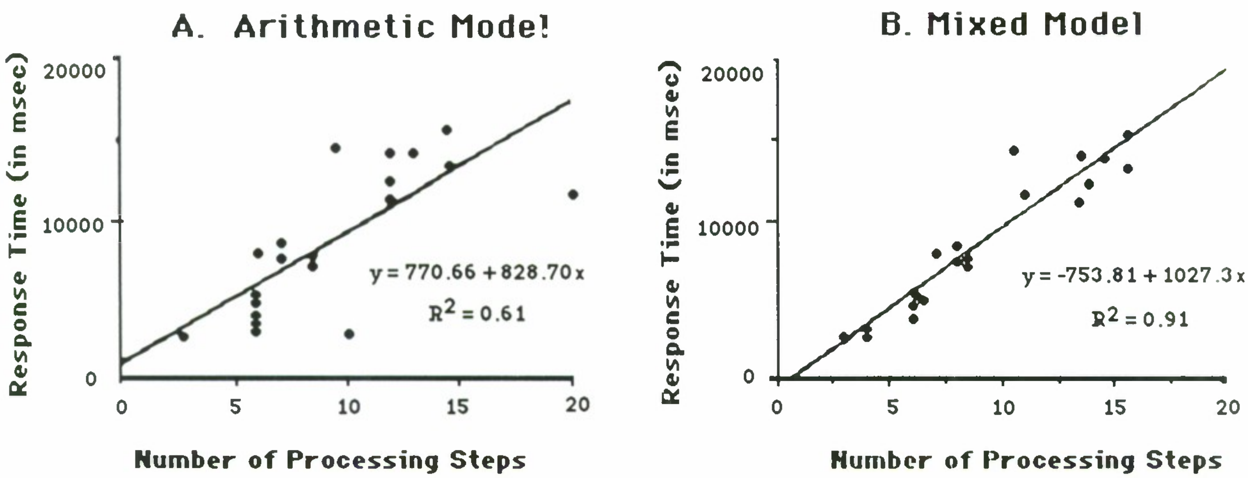


Figure 1. Response times for answering eight types of questions using three types of graphs as a function of the number of processing steps. A. Arithmetic strategy B. Mixed arithmetic-perceptual strategy.

that users vary their strategy according to the characteristics of the graph. For example, if a user were faced with a graph that had inadequate numerical labels on the axes, he or she would be forced to use the perceptual strategy to the greatest extent possible.

We have run a series of experiments to test our hypotheses about graphic processing strategies. The response time data from these experiments are consistent with a model that suggests that users tend to apply the arithmetic strategy, but will shift to the perceptual strategy under certain conditions. In the basic experiment, subjects used three types of graphs — scatter plot, a line graph, and a stacked bar graph. They were asked eight types of questions about each graph type: (1) identification — what is the value of variable A? (2) comparison — which is greater A or B? (3) addition of two numbers — A+B. (4) subtraction — A-B, (5) division — A/B, (6) mean — (A+B+C+D+E)/5, (7) addition and division by 5 — (A+B)/5, and (8) addition of three numbers A+B+C. Subjects were instructed to be as fast and accurate as possible. We predicted that the subjects' time to answer the questions using a graph would be a function of the number of processing steps required by a given strategy. Accordingly, with the arithmetic strategy, determining the mean should take longer than adding three numbers, which should take longer than adding two numbers.

We began by fitting the data to a model based on the assumption that subjects used an arithmetic strategy for all questions with all graphs. Figure 1A shows the fit of that model to the response time data. The response time generally increases as the number of processing steps increases, so the model accounts for some of the variance, 61%, but many of the data points fall far from the regression line. This model is poorest at predicting performance on two trials with the stacked bar graph — the mean and the addition of two numbers — and for the comparison trials with all three types of graphs; subjects responded on the comparison trials and the mean trial more quickly than predicted.

As discussed above, a comparison appears to be a likely task for subjects to use a perceptual strategy. In addition, the stacked bar graph intrinsically lends itself to adding the five variables by a perceptual strategy. The total height of the stack represents the cumulative value of the five variables. Accordingly, for model 2, we assumed that subjects used a

perceptual strategy to determine the cumulative value of the stacked bar graph (then looked up the value and divided by 5 arithmetically) and used only the perceptual strategy to make all comparisons. Figure 1B shows how a version of that model fits the data. This model captures a substantially greater amount of the variance, 91%, than did Model 1. In this version of the model, the regression function slope suggests that each processing step required about 1 second to complete, except for steps requiring subtraction or division (which the model assumes took 1.5 and 2 seconds, respectively).

The fit of the mixed arithmetic-perceptual model to the data, together with subjects' verbal protocols when answering questions using graphs, support our hypotheses: (1) that people use both arithmetic and perceptual strategies with graphics, (2) that for many typical questions, the bias appears to be for the arithmetic strategy (perhaps because of the greater accuracy with that strategy), and (3) subjects switch strategies as a function of the characteristics of the question and graph.

A THEORY OF GRAPHIC COMPREHENSION

The focus of the rest of this paper is on an overall theory of graphical comprehension designed to help in the development of graphic displays. The theory covers the entire process of graphic comprehension from the motivation to look at a graph, to the use of the graph, to remembering the graph.

In general, when I look at a graph, I have a particular purpose in mind — I am usually trying to answer a specific question. Thus, stage 1 in graphic comprehension would consist of either forming a representation of the question to be answered (if the question had to be remembered), or producing the question by inference or generalization. The final cognitive representation of the question would probably be much the same, regardless of whether I read it, remembered it, or generated it. The likely representational format for the question would be a semantic network (e.g., [1] and [8]). Determining the answer to the question would function as the goal of my graphic comprehension.

At the start of the second stage in graphic comprehension, I would look at the graph. On looking at the graph, I would encode the primary global features — the presence or absence of the axes and the type of graph. These would be encoded in a format that would permit reproduction of certain lower level features, such as the orientation of both the elements that make up the graph type and the axes. For example, subjects in our representation experiments generally recalled the horizontal orientation of the bars in a column graph, despite (or, perhaps, because of) their difference from the more typical vertical bar graphs. Interestingly, features that one might expect to be important to a graph user, such as the number of data points, appear not to be encoded as part of this global encoding stage. One hypothesis of this model is that features represented during the global encoding stage receive the bulk of the representational strength. That is to say, they will be the best remembered.

The third stage in graphic comprehension is to use the goal and the global features of the graph to select a processing strategy. If my goal were to compare the value of variables or (possibly) to compare a trend, I would select a perceptual strategy. If my goal were to determine the sum of four variables, and numbered axes were present and the graph type supported it (e.g., a line graph or a bar graph), then I would select the arithmetic strategy.

During the next stage, I would implement the processing steps called for in the strategy determined in the third stage. For example, adding variables A and B from a line graph would involve the following processing steps: (1) locate the name of variable A on the X axis, (2) locate variable A in the x-y coordinate space of the body of the graph, (3) locate the value of variable A on the Y axis and store in working memory, (4) locate the name of variable B on the X axis, (5) locate variable B in the x-y coordinate space of the body of the graph, (6) locate the value of variable B on the Y axis and store in working memory, and (7)

add the value of variable A to the value of variable B to produce the value "sum."

Because the semantic and quantitative information (i.e., the variable names and values, respectively) are processed to some extent during this phase, some of that information will be represented, but, as our recall data suggest, not strongly. As a final stage in graphic comprehension, I would examine the result from processing step 7, the "sum," to determine if it plausibly met the goal set in comprehension stage 1. If the response was a plausible fit with the goal, I would incorporate the answer into the semantic network that represented the goal.

This theory directs both future research in graphics and the design of graphical computer interfaces. For example, future research will be needed to determine specific processing models for different questions using the perceptual strategy. In addition, predictions about the memory for quantitative and semantic information in a graph need to be tested. Finally, many of the design principles derived from the theory are concerned with the complex relations between the task (or goal), the characteristics of the graphical display, and the processing strategies. For example, if a subject is likely to use arithmetic strategy (e.g., with an addition or subtraction question), the axes should be numbered with sufficient numerical resolution. The graph type should allow the user to read a variable's value directly from the axis and should not require multiple computations to determine a variable's value (as a stacked bar graph does). One of our long-term goals is to produce a model of graphic comprehension that is sufficiently elaborate to allow us to build tools to aid in the design of graphical interfaces.

ACKNOWLEDGEMENTS

The authors would like to thank Sarah Breedin and Tim McKay for helpful comments on this paper and, along with Kritina Holden and Susan Adam, for helpful discussions about graphics.

REFERENCES

1. Anderson, J. R. *The Architecture of Cognition*, Harvard University Press, Cambridge, MA, 1983.

2. Bertin, J. *Seminology of Graphics* (translated by W. J. Berg), University of Wisconsin Press, Madison, WI, 1983.

3. Card, S. K., Moran, T. P., and Newell, A. *The Psychology of Human-Computer Interaction*, L. Erlbaum Associates, Hillsdale, NJ, 1983.

4. Cleveland, W. S., and McGill, R. "Graphical perception: Theory, experimentation, and application to the development of graphical methods," *Journal of the American Statistical Association*, 79, 1984, 531-554.

5. Cleveland, W. S., and McGill, R. "Graphical perception and graphical methods for analyzing and presenting scientific data," *Science*, 229, 1985, 828-833.

6. Gillan, D. J., Lewis, R., and Rudisill, M. "Models of user interaction with graphical interfaces: I. Statistical graphs," *Proceedings of CHI*, 1989, ACM, New York, 375-380.

7. Hutchins, E. L., Hollan, J. D., and Norman, D. A. "Direct manipulation interfaces," *User-System Centered Design*, L. Erlbaum Associates, Hillsdale, NJ, 1986, 87-124.

8. Kintsch, W., and van Dijk, T. A. "Toward a model of text comprehension and production," *Psychological Review*, 1978, 85, 363-394.

9. Shepherd, R. N., Romney, A. K., and Nerlove, S. B. *Multidimensional Scaling: Theory and Application in the Behavioral Sciences*, Seminar Press, New York, 1972.

10. Tufte, E. R. *The Visual Display of Quantitative Information, Graphics Press*, Cheshire, CT, 1983.

DESIGNERS' MODELS OF THE HUMAN-COMPUTER INTERFACE

Douglas J. Gillan
Rice University

Sarah D. Breedin
Lockheed Engineering and Sciences Company

Research directed by Marianne Rudisill, Manager, Human Computer Interaction Lab, NASA JSC.

INTRODUCTION

Do people's cognitive models of the human-computer interface (HCI) differ as a function of their experience with HCI design? A cognitive model can be defined as a representation of a person's knowledge consisting of (1) a set of elemental concepts (elements in a model of an HCI might include windows, menus, tables, and graphics), (2) the relations among the elements (for example, a mouse and a touch screen might be related as input devices), and (3) the relations among groups of associated elements (for example, a group of input devices might be related to a group of user-computer dialogue techniques). (See [4], [7], and [10] for additional definitions.)

Cognitive modeling in the area of human-computer interaction has generally focused on how the user represents a system or a task [4]. The results of this approach provide information relevant to Norman's concept of a user's model [9]. In contrast, the present paper focuses on the models of HCI designers, specifically on designers' declarative knowledge about the HCI. Declarative knowledge involves the facts about a given domain and the semantic relations among those facts (e.g., [1]); for example, knowing that the mouse, trackball, and touch screen are all types of interactive devices. The results of our approach provide information relevant to Norman's concept of a design model [9].

Understanding design models of the HCI may produce two types of benefits. First, interface development often requires inputs from two different types of experts—human factors specialists and software developers. The primary work of the human factors specialists may involve identifying the ways in which a system should display information to the user, the interactive dialogue between the user and system, and the types of inputs that the user should provide to the system. The primary work of the software developers may center around writing the code for a user interface design and integrating that code with the rest of the system. Given the differences in their backgrounds and roles, human factors specialists and software developers may have different cognitive models of the HCI. Yet, they have to communicate about the interface as part of the design process. If they have different models, their interactions are likely to involve a certain amount of mis-communication. Second, the design process in general is likely to be guided by designers' cognitive models of the HCI, as well as by their knowledge of the user, tasks, and system. Designers in any field do not start with a *tabula rasa*; rather they begin the design process with a general model of the object that they are designing, whether it be a bridge, a house, or an HCI.

Our approach to a design model of the HCI was to have three groups make judgments of categorical similarity about the components of an interface: (1) human factors specialists with HCI design experience, (2) software developers with HCI design experience, and (3) a baseline group of computer users who had no experience in HCI design. The components of the user interface included both display components such as windows, text, and graphics, and user interaction concepts, such as command language, editing, and help. The judgments of the three groups were analyzed using hierarchical cluster analysis [8],

and Pathfinder ([12] and [13]). These methods indicated, respectively, (1) how the groups categorized the concepts, and (2) network representations of the concepts for each group. The Pathfinder analysis provides greater information about local, pairwise relations among concepts, whereas the cluster analysis shows global, categorical relations to a greater extent.

METHOD

SUBJECTS

Thirty-five subjects (members of a NASA Space Station *Freedom* user interface working group, employees at Lockheed and AT&T Bell Laboratories, and students at Rice University) were assigned to one of three groups on the basis of their work and/or academic experience in human factors and software development: human factors specialists (n = 13), software developers (n = 11), and computer users with no experience in HCI design (n = 11). The human factors specialists reported that their median years of working experience in human factors was 4.5, in user interface issues was 4.5, and in software development was 2. The software development group reported substantially more software experience than the human factors group, a median of 5.5 years of work, slightly longer experience with interface issues, a median of 6 years, but markedly less human factors experience, 1 year. The non-HCI group's relevant experience was minimal, with only software courses (median number of courses = 1) and experience as users of software (primarily for word processing).

MATERIALS

A questionnaire was designed to investigate individual's models and knowledge of the HCI. The first part of the questionnaire consisted of a list of 50 HCI terms (for example, auditory interface, characters, command language, and keystroke) selected from (1) the indices of CHI Proceedings from 1986 to 1988 and (2) recent general books on human-computer interaction ([2], [3], [10], and [11]). Terms were selected based, in part, on their co-occurrence in these sources and the frequency of occurrence within the sources. The terms were presented in alphabetical order.

The final part of the questionnaire asked for information about the subject's experience with and knowledge of human factors and software design. The answers from this section were used in assigning subjects to one of the three groups.

PROCEDURE

Subjects read a set of general instructions that oriented them to the tasks. Included in these instructions was a comprehensive example that had the subjects apply the procedure to a set of food concepts. Then, subjects started with Part I by reading through the entire list of 50 terms.

If a subject was unfamiliar with a term, he or she was instructed to cross that term off the list. Next, subjects sorted related terms into 'piles' by writing the terms into columns on a data sheet. Subjects could place items in more than one pile or leave items out of any pile.

RESULTS

The results from Part I of the questionnaire were analyzed using two multivariate statistical techniques—hierarchical cluster analysis [8] and Pathfinder analysis ([12] and [13]). The cluster analysis indicates how subjects categorize concepts, whereas the Pathfinder analysis provides a network representation of the concepts.

CLUSTER ANALYSIS: CATEGORIES OF DECLARATIVE KNOWLEDGE

To prepare the data for the cluster analysis, a co-occurrence matrix of the concepts was created for each subject. When a subject placed two concepts in the same pile, a count was entered into the corresponding cell of the matrix. Then, the matrices for all of the subjects within a group were combined. The co-occurrence matrices for each group were converted to dissimilarity matrices by subtracting the co-occurrence value from the

number of subjects plus 1, and a minimum-distance hierarchical cluster analysis was performed.

The cluster analysis displayed in Figures 1A, 1B, and 1C shows substantial differences between the non-HCI group and the experts, but reveals some similarities and dissimilarities between the two expert groups. The data displayed includes only those clusters in which 50% or more of the subjects in that group sorted the items into the same pile. The figures show (1) subclusters with a relatively small number of concepts and for which agreement of categorical co-occurrence was the greatest, and (2) various levels of supraclusters consisting of one or more subclusters and additional concepts. The strength of agreement within a group (i.e., the percentage of subjects who placed the concepts in the same category) is indicated by the percentage in the cluster boundary and by the width of the line around a cluster (thicker lines indicate greater agreement). The label for a cluster, selected by the authors, is in bold above the cluster.

The two expert groups had both a greater number of clusters and generally more complex hierarchical relations among the clusters than did the non-HCI group. In addition, both expert groups differed substantially from the nonexpert group in the content of their clusters, with two exceptions: (1) All three groups had relatively high agreement that the terms, *expert user*[1] and *novice user*, belonged to the same cluster, which was hierarchically unrelated to other clusters, and (2) the three groups of subjects categorized *mouse*, *touch screen*, *trackball* and *interactive devices* together. However, the types of devices were not part of a larger hierarchy for the non-HCI group, but were included in the Interaction Techniques supracluster for both expert groups. Other areas of basic agreement between the two expert groups were a Guidance/Help supracluster and an Output cluster.

The cluster analysis shows two key areas of disagreement between the human factors and software experts: (1) the contents and organization of the Display Elements cluster and (2) the relation of software concepts to other user interface concepts. In the Display Elements cluster, human factors experts had three categories at the same level in the hierarchy—Textual Elements[2], Graphical Elements, and Tabular Elements. In contrast, software experts had a Graphical Elements subcluster which was nested in a Coding/Graphics subcluster, which, in turn, was nested in a larger Nontextual Display Elements subcluster. Note also that the software developers grouped *color coding* and *highlighting* in the Display Elements subcluster, whereas the human factors specialists grouped those two concepts with *data grouping* and *symbolic codes* in a separate cluster, Display Coding. This difference in categorizing display coding concepts may be due to a greater emphasis by human factors experts on the similarities in function among methods for coding information on a display.

As Figure 1B shows, the software group included six software concepts concerned with the user interface and applications in the User Interface Elements supracluster. In contrast, the human factors group categorized the software-related concepts in a separate supracluster unconnected to other user interface concepts. This finding suggests that, in the software developers' design model, software is more fully integrated with other HCI concepts than it is in the human factors specialists' model.

PATHFINDER: NETWORKS OF DECLARATIVE KNOWLEDGE

The similarity matrices derived from the sorting data for each group were also analyzed with the Pathfinder algorithm using the Minkowski r-metric, $r = \infty$ and $q = 49$ (see [13]). The Pathfinder algorithm generated a network solution for each of the three matrices. However, the network for the non-HCI group was exceedingly complex and difficult to

[1] In the description that follows, the terms from the questionnaire are italicized.

[2] Names for the subclusters are indicated in Figure 1 by a boxed label with an arrow pointing to the specified subcluster.

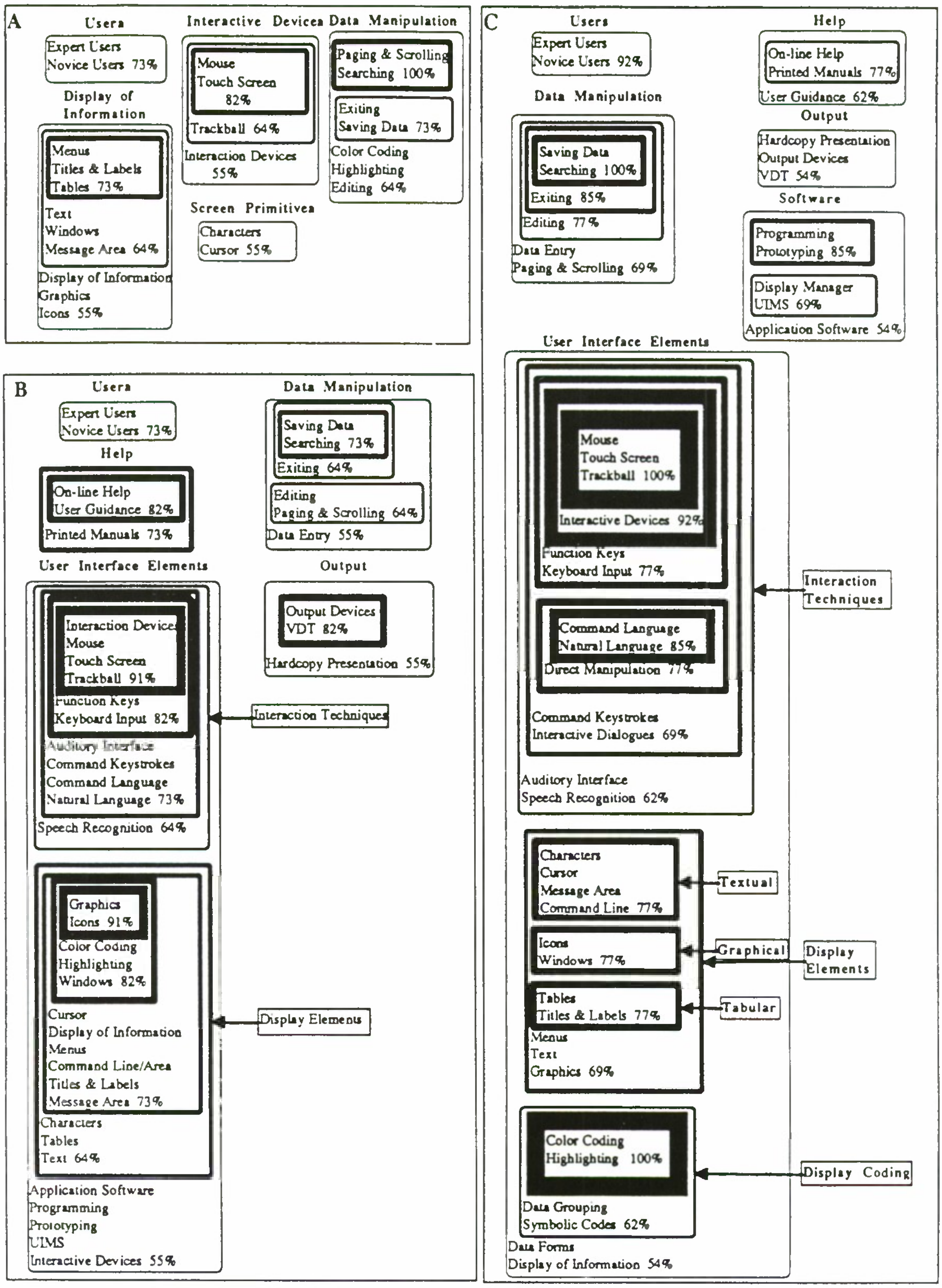


Figure 1. Cluster analysis for non-HCI subjects (1A), software developers (1B), and human factors specialists (1C).

interpret, with 171 links among the 50 concepts. Consequently, we will only focus on the more interpretable results from the two expert groups. The human factors group had 81 links and the software experts 69 links among the 50 nodes. Figure 2 shows the results of the Pathfinder analysis for the human factors (2A) and software experts (2B). The graphs show each concept as a node in a network and show the links between the nodes. The strength of each link is represented by its width, with wider lines indicating stronger connections.

Human Factors Specialists. The network representation for the human factors experts consists primarily of subnetworks of interconnected concepts, indicated in Figure 2A by the dashed lines around the groups of concepts (with subnetwork labels, selected by the authors, contained in the boxes pointing to the relevant subnetwork). Subnetworks were defined as groups of three or more concepts, in which each concept linked directly to at least two other concepts in that subnetwork, and in which the interconcept distance was no greater than two links for all concepts. This definition maintains a high level of interconnection and close association of concepts within the subnetwork. With the exception of *speech recognition*, which appears in both Input Devices and Advanced User Interface Techniques, the subnetworks are cleanly separated, in that the concepts are not shared by subnetworks.

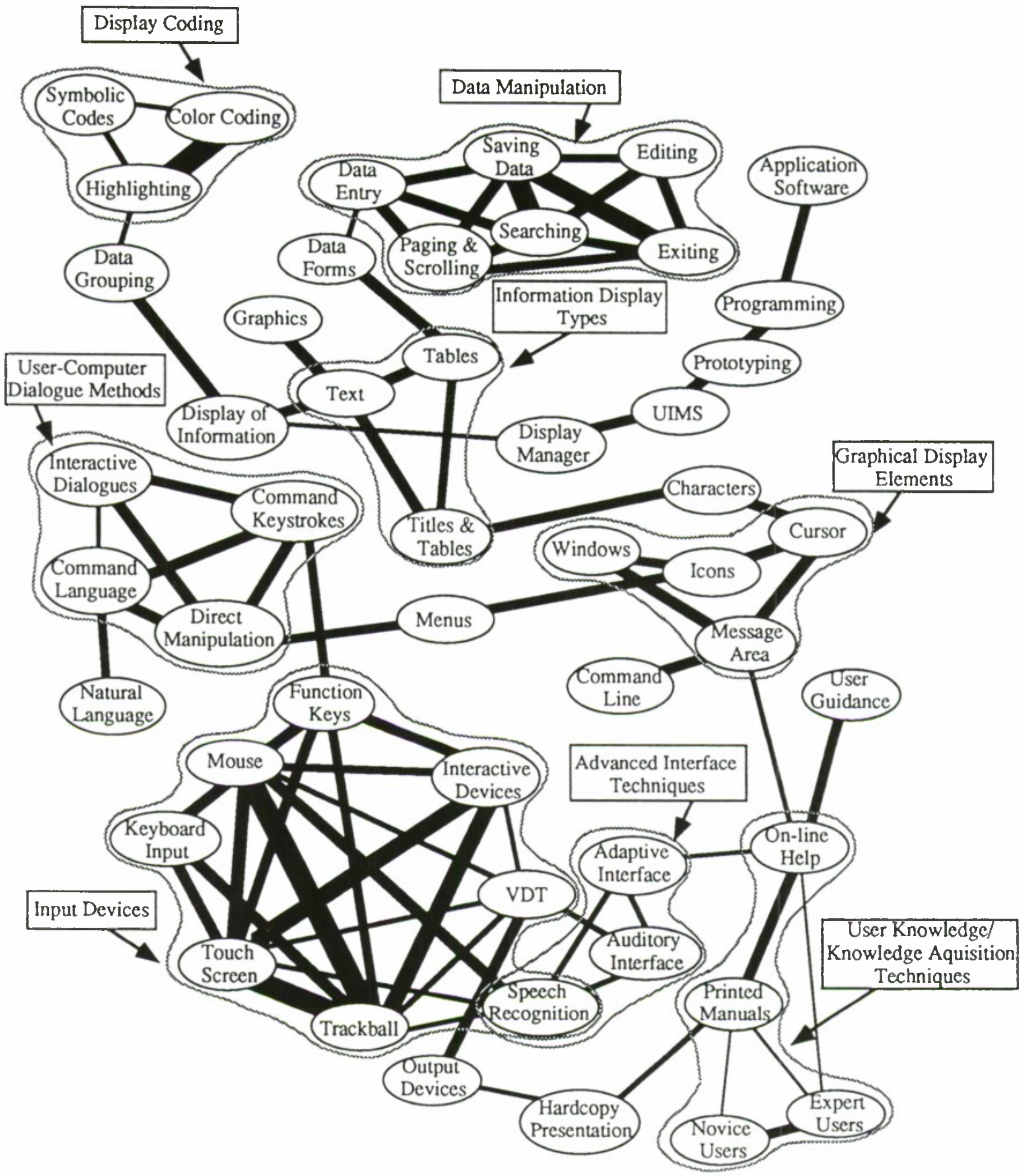
Each subnetwork for the human factors experts connects with other subnetworks. Several of the subnetworks have a direct link between two concepts. For example, the User-Computer Dialogue Methods subnetwork and the Input Devices subnetwork are connected by a link between *command keystrokes* and *function keys*. The other subnetworks make connections through one or two intermediate concepts. For example, *menus* provides a conceptual connection between User-Computer Dialogue Methods and Graphical Display Elements. Similarly, *data forms* links the Data Manipulation subnetwork to Information Display Types.

Only a few concepts are offshoots of a subnetwork unconnected to another concept—*graphics*, *natural language*, *command line*, and *user guidance*. The major departure from the subnetwork structure is the string of concepts related to software, with *display of information* linked to *display manager*, which connects with *UIMS*, which in turn links to *prototyping*, and so on.

Software Developers. The structure of the network representation for the software experts (Figure 2B) consists of both (1) central nodes from which links radiate out in axle and spoke fashion and (2) subnetworks consisting of interconnected concepts. We defined a central node as a concept with at least three links in addition to any links it might have within a subnetwork. Central nodes are shown in grey in the figure; as in Figure 2A, subnetworks are bounded by a dashed line with labels contained in boxes.

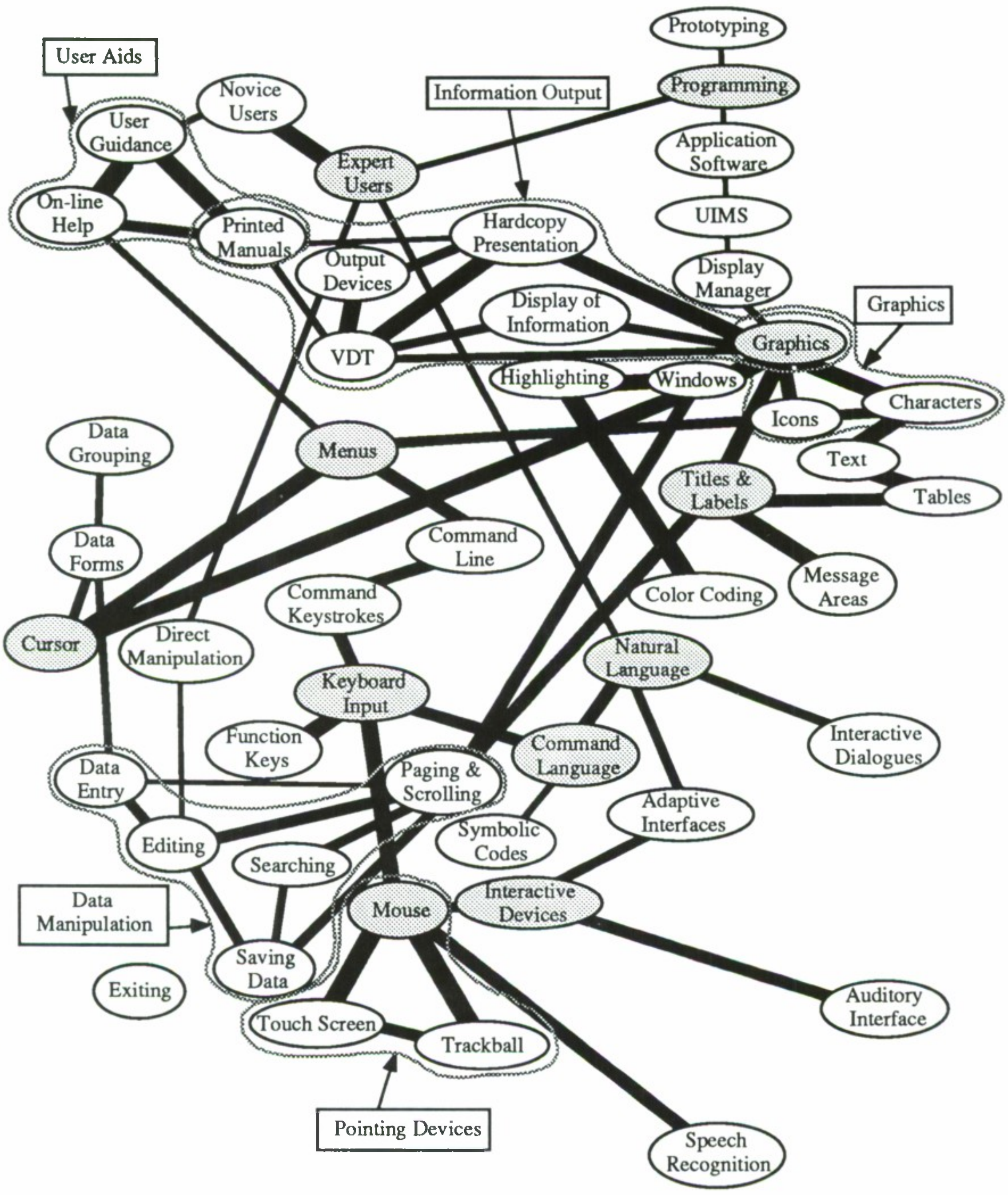
The software experts had only two subnetworks containing more than three concepts, Data Manipulation and Information Output, and had only three triads of concepts. Among the central nodes, both *mouse* and *expert users* are of interest because they link directly to other central nodes, with *mouse* having strong connections to *interactive devices* and *keyboard input* and *expert users* weakly linked to *programming* and *natural language*. In addition, *mouse* functions both as a member of a subnetwork and as a central node. *Graphics* is also well connected, with membership in two subnetworks and central node status.

Comparing the Expert Groups. The networks reveal important differences between the two expert groups. Overall, the ratio of the number of links shared by the two groups to the total number of links was 0.23. Looking at specific concepts, several of the concepts that have only one link in one group's representation are strongly interconnected in the other group's network. For example, *graphics* and *natural language* are linked directly to a number of other concepts in the software experts' network, but have only one link apiece for human factors experts.



Note: Line thickness indicates the strength of the link. Thicker lines mean greater strength.

Figure 2A. Pathfinder network for human factors specialists.



Note: Line thickness indicates the strength of the link. Thicker lines mean greater strength. Central nodes are shaded in gray.

Figure 2B. Pathfinder network for software developers.

On the other hand, *function keys* is a member of the Input Devices subnetwork and connects that subnet work to the User-Computer Dialogues subnetwork for human factors experts, but links only with *keyboard input* for software experts. An additional difference is

that the networks for the software and human factors experts show no overlap between the concepts that link to only one other concept.

When a concept has the same number of direct links for the two groups, it may reveal important differences in the design models if it differs in the other concepts to which connections are made. For example, look at *user interface management system* in Figures 2A and 2B. For both groups, one of its connections is with *display manager*, indicating knowledge of the relationship between the software that manages the entire user interface and the software that writes to the screen. For human factors experts, the other connection of *UIMS* is with *prototyping*, suggesting that the prototyping capability is an important part of a UIMS for interface designers with a human factors background. However, for software experts, *UIMS* connects with *application software*, which is consistent with the software architecture of the user interface—with the UIMS interacting with the application software, as well as the display manager.

DISCUSSION

GENERAL EFFECTS OF HCI DESIGN EXPERIENCE

The data from both the cluster analysis and Pathfinder analysis show differences as an effect of expertise in human-computer interaction. Both expert groups had (1) a greater number of clusters containing more concepts and (2) more complex hierarchical structures of the clusters than did the non-HCI group. The Pathfinder solution for the non-HCI group was a mass of links between concepts with minimal differentiation. In contrast, both expert groups showed substantial and meaningful differentiation of groups of concepts within the networks. These findings indicate that training and experience with HCI design has a clear impact on the mental model of the interface. This finding, by itself, may not be surprising. However, many people outside of the field of human-computer interaction may hold contrary opinions—for example, that HCI design is simply a matter of common sense or that computer users' experience is the equivalent of HCI design experience. The present data argue against those opinions by showing the effects of user interface design experience.

EFFECTS OF SPECIFIC HCI DESIGN EXPERIENCE

Differences between the mental models of experts and novices abound (for example, see [5]). We present evidence here that experts may differ in their cognitive models as a function of their roles and experience in a common area of expertise.

The Pathfinder analyses suggest that the different types of experts differ in the overall organization of their cognitive models. Human factors experts had a network made up of distinct subnetworks, with the subnetworks tending towards heavy internal interconnection with a single connection between subnetworks. The software experts' cognitive model had multiple organizing schemes, including central nodes, as well as complex and simple subnetworks. Cooke, Durso, and Schvaneveldt [6] have shown that the network representations derived by Pathfinder are related to recall from memory, with closely linked items in the Pathfinder network being more likely to be recalled together. Consequently, recall of an HCI concept may tend to have an effect that is localized within the subnetwork for human factors experts. However, recall of that same concept may spread more broadly for software experts. For example, a software developer who thinks of *keyboard input* would be likely to recall *mouse*, *function key*, *command keystrokes*, and *command language*. In contrast, *keyboard input* would be most likely to produce recall of only *mouse* and *function keys* for human factors experts. The localization of recall might help human factors experts to maintain a more focused stream of thought, but the broader spread of recall may help software experts to think more innovatively about HCI concepts by activating more varied concepts.

Differences in the concepts that are linked or in the categories in which HCI designers place

concepts might be expected as a function of experience. For example, software developers would be much more likely to see the relations between software and other HCI concepts than would human factors specialists. However, why would these two groups have very different organizing schemes for their concepts? One possibility is that software developers have to be concerned with both the ways in which the HCI software will be used and with the methods for implementing the software. In other words, their cognitive model may represent a compromise between knowledge about the function and about the implementation of the human-computer interface. In contrast, the cognitive model of human factors specialists may be more closely tied only to function.

PRACTICAL IMPLICATIONS

The Pathfinder and cluster analyses showed substantial differences in the number of connections and the conceptual links for a variety of the HCI concepts, such as *graphics* and *function keys*. These findings suggest that design team members with different types of expertise should take care to define their terms when discussing the conceptual categories—user interface elements and display coding—and about specific concepts like graphics, function keys, speech recognition, and natural language. A term like graphics may evoke a more elaborate set of associated concepts for design team members with backgrounds in software development than it does for those in human factors, whereas function key may evoke more concepts for human factors specialists.

One way of eliminating the problems of miscommunication due to different design models might be to train all of the designers to think alike. However, even if this were possible, it might lead to unintended problems in user interface design. Diversity of thinking may improve the design process. Thus, training out the diversity might result in a team that could not make conceptual breakthroughs or recognize when they were going down a blind alley. The best user interface designs are likely to emerge when the human factors specialists on a team can think their way and the software developers can think their way, but when each member understands the meaning of the others' thoughts when expressed in language or design. The representation of design team members' cognitive models described in this paper provides the first step in enhancing that understanding.

ACKNOWLEDGEMENTS

This research was funded by NASA Contract NAS9-17900 and was performed within NASA Johnson Space Center, Human-Computer Interaction Laboratory. The authors would like to thank Dr. Nancy I. Cooke for her advice on the analysis methods used in this paper and her comments on the paper itself.

REFERENCES

1. Anderson, J. R. (1983). Architecture of cognition. Cambridge, MA: Harvard University Press.

2. Baecker, R. M., and Buxton, W. A. S. (1987). *Readings in human-computer interaction; A multidisciplinary approach*. Los Alamos, CA: Morgan Kaufmann Publishers, Inc.

3. Card, S. K., Moran, T. P., and Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

4. Carroll, J. M., and Olson, J. R. (1988). Mental models in human-computer interaction. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 45-65). Amsterdam: Elsevier Science Publishers.

5. Chi, M. T. H., Feltovich, P. J., and Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.

6. Cooke, N. M., Durso, F. T., and Schvaneveldt, R. W. (1986). Recall

and measures of memory organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 538-549.

7. Gentner, D., and Stevens, A. (Eds.). (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

8. Miller, G. A. (1969). A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology*, 6, 169-191.

9. Norman, D. A. (1986). Cognitive engineering. In D.A. Norman & S. Draper (Eds.), *User-centered system design. New perspectives in human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

10. Norman, D. A., and Draper, S. (1986). *User-centered system design: New perspectives in human-computer interaction* (pp. 31-61). Hillsdale, NJ: Lawrence Erlbaum Associates.

11. Shneiderman, B. (1987). *Designing the user interface: Strategies for effective human-computer communication*. Reading, MA: Addison-Wesley.

12. Schvaneveldt, R. W., and Durso, F. T. (1981). Generalized semantic networks. Paper presented at the *Meetings of the Psychonomics Society*.

13. Schvaneveldt, R. W., Durso, F. T., and Dearholt, D. W. (1985). Pathfinder: Scaling with network structures. *Memorandum in Computer and Cognitive Science, MCCS-85-9*, Computing Research Laboratory, New Mexico State University.

AUTOMATED SYSTEM FUNCTION ALLOCATION AND DISPLAY FORMAT: TASK INFORMATION PROCESSING REQUIREMENTS

Mary P. Czerwinski
Lockheed Engineering and Sciences Company

Research directed by Marianne Rudisill, Manager, Human Computer Interaction Lab, NASA JSC.

INTRODUCTION

Questions relevant to the Human Factors community attempting to design the display of information presented by an intelligent system are many: What information does the user need? What does the user have to do with the data? What functions should be allocated to the machine versus the user? Currently, Johnson Space Center is the test site for an intelligent Thermal Control System (TCS), TEXSYS, being tested for use with Space Station Freedom. The implementation of TEXSYS' user interface provided the Human-Computer Interaction Laboratory with an opportunity to investigate some of the perceptual and cognitive issues underlying a human's interaction with an intelligent system.

An important consideration when designing the interface to an intelligent system concerns function allocation between the system and the user. The display of information could be held constant, or "fixed," leaving the user with the task of searching through all of the available information, integrating it, and classifying the data into a known system state. On the other hand, the system, based on its own intelligent diagnosis, could display only relevant information in order to reduce the user's search set. The user would still be left the task of perceiving and integrating the data and classifying it into the appropriate system state. Finally, the system could display the *patterns* of data. In this scenario, the task of integrating the data is carried out by the system, and the user's information processing load is reduced, leaving only the tasks of perception and classification of the patterns of data. Humans are especially adept at this form of display processing [1, 2, 11, and 12].

Although others have examined the relative effectiveness of alphanumeric and graphical display formats [7], it is interesting to reexamine this issue together with the function allocation problem. Expert TCS engineers, as well as novices, were asked to classify several displays of TEXSYS data into various system states (including nominal and anomalous states). Three different display formats were used: *fixed* (the TEXSYS "System Status at a Glance"), *subset* (a relevant subset of the TEXSYS "System Status at a Glance"), and *graphical*. These three formats were chosen due to previous research showing the relevant advantages and disadvantages of graphical versus alphanumeric displays (see Sanderson et al., 1989 for a review), and because of the vast amount of literature on the beneficial effects of reducing display size during visual search in cognitive psychology (see Shiffrin and Schneider, 1977; Schneider and Shiffrin, 1977). The hypothesis tested was that the graphical displays would provide for fewer errors and faster classification times by both experts and novices, regardless of the *kind* of system state represented within the display [11]. The subset displays were hypothesized to be the second most effective display format/function allocation condition, based on the fact that the search set is reduced in these displays [5, 6]. Both the subset and the graphic display conditions were hypothesized to be processed more efficiently than the fixed display condition, which corresponds to the "System Status at a Glance" display currently used in TEXSYS.

METHOD

SUBJECTS

Four frequent users of TEXSYS, thermal control engineers at JSC, participated in the experiment. The subjects had an average of

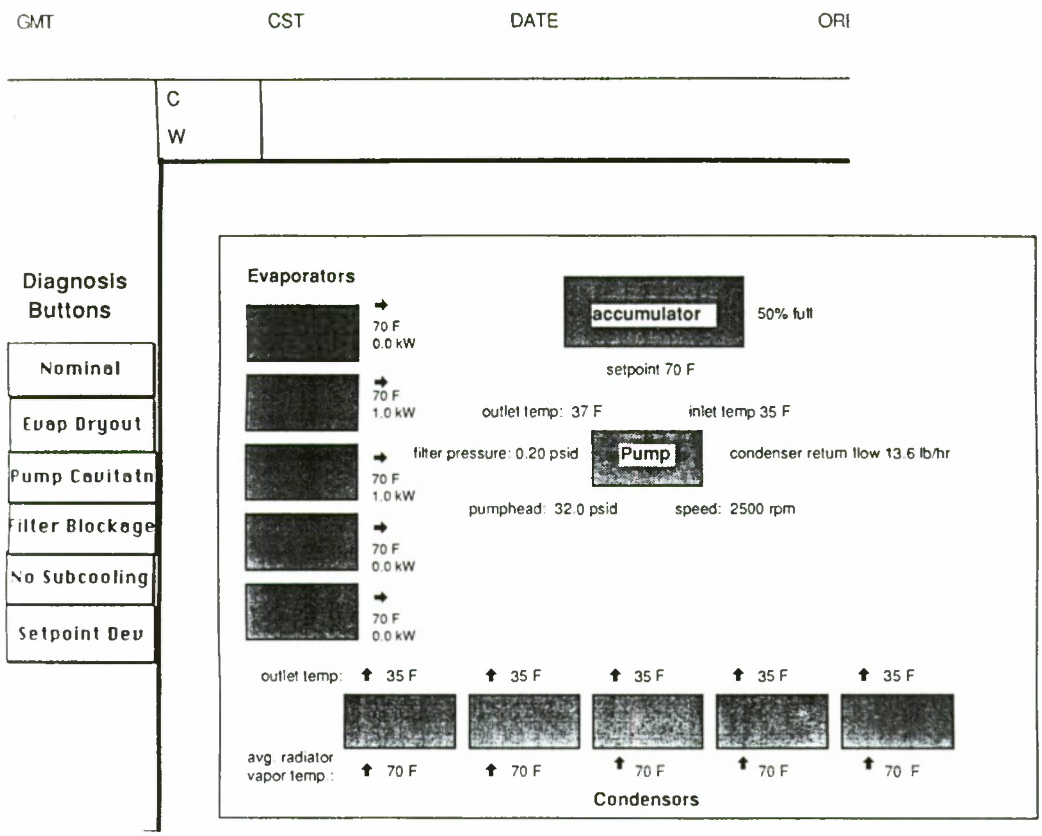


Figure 1. The "fixed" display.

eight years experience. Six novices, all engineers, also participated in the experiment. None of the novice subjects was familiar with the two-phase thermal bus system used in the TEXSYS project, nor with thermal control systems in general. All subjects were experienced users of Macintosh computers, and all had normal or corrected-to-normal vision.

STIMULI AND MATERIALS

The design, presentation, and collection of all stimulus materials and data were carried out on a Macintosh IIx computer using SuperCard and SuperTalk. A mouse was used for all subject inputs. Examples of the fixed, subset, and graphical display formats can be seen in Figures 1, 2, and 3, respectively. Note that, while the fixed and graphical displays both contain information about all of the major system components, the subset displays only show a subset of the system data.

System Faults. Five different system anomalies could occur during the experiment: evaporator dryout, filter blockage, pump cavitation, loss of subcooling and setpoint deviation.

MATCHING NOMINAL AND ANOMALOUS DISPLAYS

Nominal displays were matched with anomalous displays for two reasons. First, designing the experiment in this manner avoids biasing the subjects toward responding "fault" or "no fault." The second reason is related to a peculiarity in the subset display condition. In these displays, subjects were told that the expert system had made a reasonable guess as to the critical system state, and only information concerning that state was shown. In nominal conditions, in order to control for the amount of information displayed to the subject, the same component subsets were shown as in the fault conditions. However,

Figure 2. The "adaptive" display.

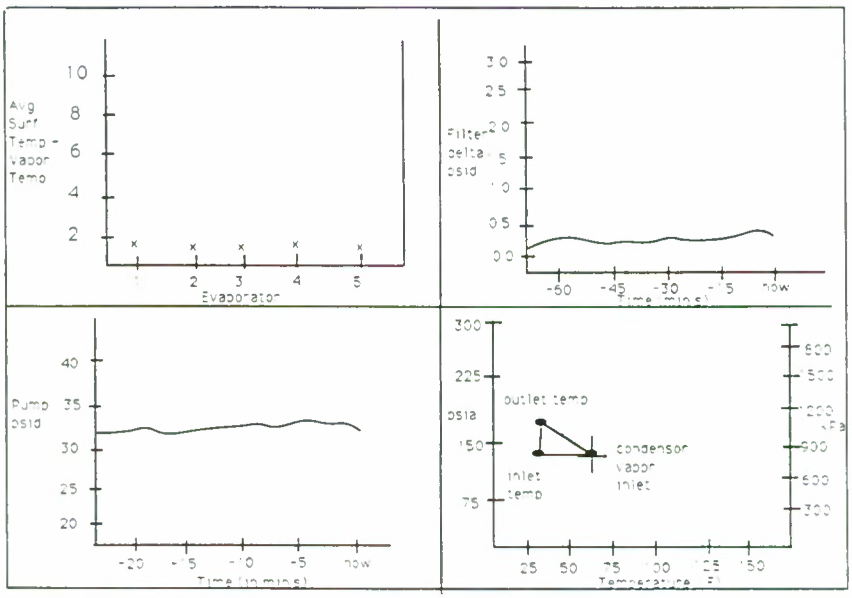


Figure 3. The "graphic" display.

since the displays were nominal, the displayed data values were never aberrant. The matching of displays simply involved replicating the no-fault displays and then changing particular component values to off-nominal for the fault displays.

DESIGN

The experimental design was a 3 x 2 x 5 x 2 factorial, with three different display formats (fixed, subset, and graphic), both nominal and anomalous display instances, five different state instances, and two repetitions per condition. Note that this design implies that a system fault occurred on 50% of the trials. There were two groups of subjects run in the experiment: experts and novices. The novices were given two sessions of training, which added an extra factor (session) to their design. All variables were run within subjects, but experts and novices were analyzed separately. The three different display formats were blocked, such that there were three blocks of 20 trials (including the repetitions) in each experimental session. The order in which each subject received the three display formats was counterbalanced. All of the other factors were randomized within a display condition block. The dependent measures collected were reaction time and percent correct.

PROCEDURES

Experts. During an orientation, prior to actual data collection, the experts were shown a table of nominal data values (as well as the acceptable ranges of deviation for those values) for the major components of the system.

Novices. The same materials that were used for orientation of the experts were used to train the novices. Unlike the expert subjects, the novices studied the nominal operations table for approximately 50 minutes[1]. During this time, they were informed about the patterns of data which might occur for each of the five system faults[2].

Both expert and novice subjects were instructed to monitor the displays presented to them for one of the six system states. They were instructed to search the system display quickly, without making errors, for system status information. Once the displayed data had been categorized by the subject, s/he was instructed to indicate which system state had occurred via a button-click with the mouse input device.

All subjects were run through a practice experiment, in which an example of each Display Format x System State combination was included. Feedback in the case of an error was provided for the subjects as a computer beep.

The diagnosis buttons were located to the far left of the display, as can be seen in Figure 1. The CONTINUE button (on the intertrial screen) was located in the center of the position previously occupied by the six diagnosis buttons. This button placement was used in order to reduce the motor movement time involved in selecting any of the six diagnosis buttons. Trials were self-paced, and subjects were encouraged to take a short break between blocks. The experimental session lasted approximately one hour.

RESULTS AND DISCUSSION

ERRORS

Experts. Overall, the experts operated at an accuracy level of 93% correct. A separate analysis of variance (ANOVA) with repeated measures was run on the error data for both

---

[1]This was the average amount of time needed to train each individual subject, although each subject's time varied slightly due to the number of questions they asked.

[2]Novice subjects were run through the experiment for two reasons: there were too few experts available to participate in the experiment, and the experts were extremely well-practiced at diagnosing the System Status-at-Glance displays. Both problems might have biased results. The extra novice session was to ensure that novice subjects had a chance to attain near-expert levels of performance in this task.

TABLE 1.

Average Logged Reaction Times for Diagnosing the Six System States in Each Display Format for Expert and Novice Subjects.

	Experts			Novices		
State	Fixed	Adaptive	Graphic	Fixed	Adaptive	Graphic
Nominal	9.3	8.5	9.4	8.5	8.0	8.1
Evap Dryout[3]	9.4	9.1	9.8	8.3	7.6	7.9
Filter Block[4]	9.4	8.7	8.9	8.9	8.3	8.0
Pump Cav[5]	9.1	8.6	8.8	8.5	8.0	7.7
No Subcooling[6]	9.5	9.3	9.7	8.5	8.0	8.6
Setpoint Dev[7]	9.3	8.2	8.9	8.9	8.6	8.7
Average	9.3	8.7	9.3	8.6	8.1	8.2

experts and novices. For experts, the ANOVA was a 3 x 2 x 5 x 2, representing the factors of display (fixed, subset, and graphic), fault or no fault, type of fault, and repetition. The analysis revealed a significantly larger number of errors with nominal displays, $F(1,3) = 22.09$, $p < .02$. No other effects were significant for the experts.

Novices. On the average, the novice subjects performed at an accuracy level of 91.2% correct in session 1, and 93% correct during session 2. For novice subjects, a 2 x 3 x 2 x 5 x 2 ANOVA with repeated measures was carried out on the error data. The first variable corresponds to the two sessions of training that novice subjects received during the experiment; all other factors are identical to those used in the expert subject's ANOVA. There was a significantly larger number of errors in the nominal display condition, $F(1,5) = 20.05$, $p < .01$. No other effects were significant.

REACTION TIMES

A *t*-test was performed between the overall average reaction times of the experts and the overall average (across two sessions) of the novices. No significant difference was found between the two groups[8], $t(8) = 1.61$, $p > .05$.

Experts. The pattern of results for the expert subjects can be seen in Table 1. The ANOVA revealed significant main effects of display condition, $F(2,6) = 7.9$, $p < .05$, with subset displays processed the most quickly, followed by the graphical displays. No other main effects were significant for the expert subjects. However, there was a significant interaction between whether or not a fault was present and which *type* of fault had to be diagnosed, $F(4,12) = 3.27$, $p < .05$. This interaction reflected the fact that there were larger response time differences within the anomalous display instances than within the nominal displays, although planned comparisons did not reveal any significant differences between the anomalous display instances (all p's > .05).

---

[3]Evaporator Dryout

[4]Filter Blockage

[5]Pump Cavitation

[6]Loss of Subcooling

[7]Setpoint Deviation

[8]No significant difference was found in the error data, as well.

Novices. The pattern of results for the novice subjects is shown in Table 1. The ANOVA revealed significant main effects of session, $F(1,5) = 38.33$, $p < .01$; display condition, $F(2,10) = 14.04$, $p < .01$; and type of fault being diagnosed, $F(4,20) = 13.51$, $p < .001$. Session 2 was faster than session 1, and, again, the subset displays were processed most quickly. A significant interaction occurred between display condition and the *type* of fault being diagnosed, $F(8, 40) = 2.76$, $p < .05$. This interaction was not observed for the expert subjects, and reveals a pattern of data whereby certain faults are processed more quickly in particular formats. Finally, there was a significant interaction between whether or not a fault was occurring and the type of fault to be diagnosed, $F(4,20) = 3.98$, $p < .05$. This interaction is similar to that observed in the expert data. This interaction reflected the fact that, for nominal conditions, none of the display instances were processed significantly faster than the average of the others, as determined by planned comparisons (all p's > .05). However, in the fault condition, the evaporator dryout fault was processed significantly faster than the average of the other faults, $t(9) = -1.88$, $p < .05$, and the setpoint deviation fault was processed significantly slower than the average of the other faults, $t(9) = 2.13$, $p < .05$.

Finally, it should be noted that for both the experts and the novices there was probably a speed-accuracy trade-off operating on the reaction times within the no-fault condition. Specifically, errors increased significantly in the nominal condition, while reaction times were no different than those in the fault displays. This may have masked any significant effects occurring in the no-fault display conditions.

EXPERIMENT 2

Experiment 1 demonstrated the benefit of showing only relevant information to the subject. It was also shown that novices appear to diagnose certain faults better in a subset, alphanumeric format, while other fault diagnoses benefit from a graphical display format. However, one problem with interpreting this result has to do with the fact that the amount of information was not controlled between the subset alphanumeric and the graphical display conditions. In other words, there was no subset, graphic display condition. Experiment 2 equated more fully the two conditions and it was a means by which to explore the issue that a graphical format would always be a better representation when only the relevant state information is displayed.

It was also hypothesized in Experiment 2 that the kind of information processing required while diagnosing a display could affect performance. This was because one subset of the Experiment 1 faults (evaporator dryout and loss of subcooling) could be described as requiring a serial scan of the data followed by one memory comparison in all of the format conditions (the one memory comparison refers to the comparison of the displayed data value with a memorized nominal value for that system component). All other faults required the identification of one or more data values, the same sort of mental comparison with a nominal value, and then a further comparison with other component values. This extra comparison step could be argued to add load to working memory, and perhaps a graphical format is better in these conditions [11]. These ideas were tested in Experiment 2 as well.

For this experiment, one of the subset displays (relevant to the evaporator dryout fault) was used throughout the entire experiment. In one half of the experiment, subjects simply scanned evaporators to detect off-nominal surface temperatures in both graphical and alphanumeric display formats. In another half of the experiment, an extra comparison step was required in order to diagnose the data displayed in both formats.

METHOD

SUBJECTS

Seventeen Lockheed Engineering and Sciences engineers voluntarily participated in the experiment. All subjects were naive

concerning the operation of the automated Thermal Control System being simulated.

STIMULI AND MATERIALS

For the "scanning" level of the decision-making variable, the alphanumeric displays from the subset condition in Experiment 1 were used for this experiment. The graphical display was modified from Experiment 1 for this condition, so that a bar graph format was used. For the "scan + compare" condition, pump information was added to each of these display formats. Essentially, a pump outlet temperature was added to the displays for comparison with the evaporator information.

DESIGN

The experiment was a 2 x 2 x 2 factorial design, with two levels of the kind of decision-making steps required to diagnose a fault (scan, and scan + compare), both alphanumeric and graphical display formats, and nominal vs. anomalous display instances. Nested within the anomalous display instances, and only within the scan + compare conditions, was another factor — type of anomalous fault. This variable could not be added to the nomalous displays because nomalous displays do not fall into subcategories in this system. However, we did vary the particular data values within the nomalous displays so that the nomalous and anomalous displays were balanced in the number of *unique* system instances presented to any given subject during a session. This was because more faults were available for diagnosis when pump information was present in the display. Specifically, during the scan + compare trials, the subject had to distinguish four different system states: nominal, evaporator dryout, pump cavitation, or setpoint deviation. Note that in the scan only condition nominal and anomalous trials are equated, while in the scan + compare condition the subject received three times as many anomalous trials as nominal. Both the decision-making and the format variables were blocked, and the order in which subjects received the decision-making conditions was counterbalanced. However, if a subject randomly received the scan only (or scan + compare) decision-making condition first, that subject always received both display format conditions (in a random order) prior to diagnosing the scan + compare (scan only) blocks of the experiment. The magnitude and pattern of the faults within the displays were controlled across the graphic and alphanumeric display formats.

PROCEDURE

The procedure for running this experiment was identical to that for Experiment 1, although only novice subjects were run for a single session.

RESULTS AND DISCUSSION

ERRORS

The errors were submitted to an ANOVA, including the variables of decision-making steps, display format, and type of response (nominal or anomalous). There was no significant pattern of errors.

REACTION TIMES

The reaction time results are shown in Figure 4. The reaction times were submitted to an overall ANOVA, including the variables of decision-making steps, display format, and type of response (nominal or anomalous). The analysis revealed significant main effects of decision-making condition, $F(1,16) = 89.85$, $p < .001$, and display format condition, $F(1,16) = 34.72$, $p < .001$. The scanning only condition was diagnosed more quickly than the scanning and comparing condition, while the graphical format was processed more quickly than the alphanumeric display format. The interaction of decision-making condition and display format was not significant, $F(1,16) = 1.3$, $p = .2$. However, the interaction of display format condition and system state (nominal vs. anomalous) was significant, $F(1,16) = 7.37$, $p < .05$. Finally, a significant three-way interaction was observed between decision-making condition, display format, and system state, $F(1,16) = 9.16$, $p < .01$. The higher-level interactions

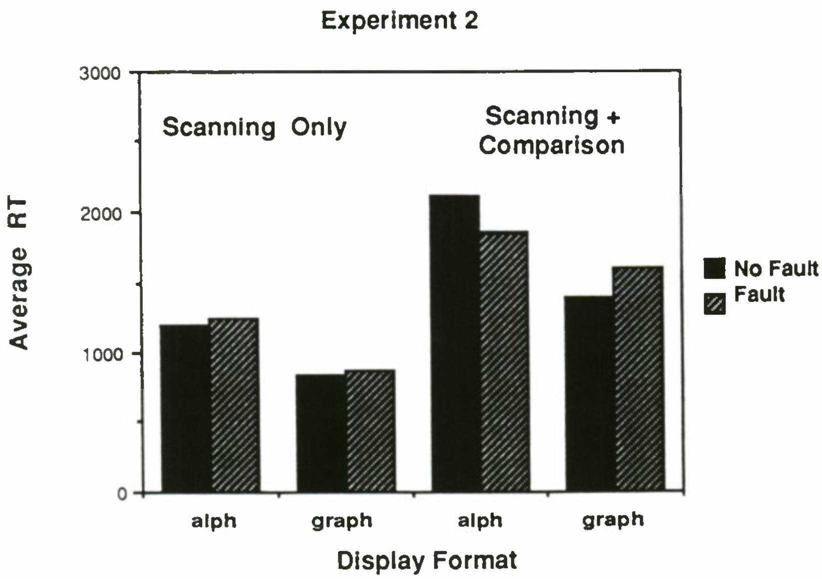


Figure 4. Average reaction time data as a function of decision-making condition and display format in Experiment 2. (alph = alphanumeric, graph = bar graph format).

reflect the fact that nominal (no fault) conditions were detected more readily than faults in all conditions except with the alphanumeric display format involving both scanning and comparing.

The results observed in Experiment 2 showed that diagnosing a subset graphical display took less time than diagnosing a subset alphanumeric display. The scanning only versus scanning and comparison manipulation could be argued to have increased the subjects' processing requirements, since diagnosis times were significantly longer in that condition. However, this increase in processing load did not lead to the interaction between display format and fault type observed in Experiment 1. It may be that the bar graph is a better way of representing data than the graphical representations used in Experiment 1. Several researchers have reported the integral processing benefits of a bar graph representation [3, 4, and 9]. Subjects may have been capitalizing on the configural [8] properties inherent in the bar graph representation in both decision-making conditions. This may be especially important when processing load is high. Some data to suggest that the bar graph representation is beneficial during heavy processing load conditions was observed in the three-way interaction reported in Experiment 2. The pattern of data showed that in the scanning and comparing condition subjects were faster at diagnosing faults in the alphanumeric displays (although still slower than in the graphic displays). Perhaps subjects were reverting to a serial search through the data in the former conditions, due to the high cognitive demands of the task. An obvious test of this notion would be to vary the number of system components showing aberrant data values for this task, in both alphanumeric and bar graph display formats. (In Experiments 1 and 2, only one system component was ever showing

off-nominal data values within a display). If subjects revert to scanning in either of the display format conditions due to heavy cognitive task demands, diagnosis times should be shorter, on the average, the greater the number of off-nominal system components [10]. This experiment is currently being run in our laboratory.

ACKNOWLEDGEMENTS

This research was funded by the National Aeronautics and Space Administration, Office of Aeronautics, Exploration and Technology (OAET), and was performed at the NASA/ Johnson Space Center Human-Computer Interaction Laboratory. The author thanks Kevin O'Brien and Kim Donner for valuable contributions during the design phase of this work. The author also thanks Nancy Cooke, Shannon Halgren and Marianne Rudisill for helpful discussions and comments on earlier drafts of this paper.

REFERENCES

1. Casey, E. J. (1986). Visual display representation of multidimensional systems: The effect of information correlation and display integrality. In Proceedings of the Human Factors Society 30th Annual Meeting (pp. 430-434). Santa Monica, CA: Human Factors Society.

2. Carswell, C. M. and Wickens, C. D. (1984). Stimulus integrality in displays of system input-output relationships: a failure detection study. In *Proceedings of the Human Factors Society*, 28th Annual Meeting.

3. Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, *68*, 361-368.

4. Coury, B. G., Boulette, M. D. and Smith, R. A. (1989). Effect of uncertainty and diagnosticity on classification of multidimensional data with integral and separable displays of system status. *Human Factors, 31*, 551-569.

5. Donderi, D. C. and Zelnicker, D. (1969). Parallel processing in visual same-different decisions. *Perception & Psychophysics*, *5*, 197-200.

6. Egeth, H., Jonides, J., and Wall, S. (1972). Parallel processing of multielement displays. *Cognitive Psychology*, *3*, 674-698.

7. Keiras, D. E. (1988). Diagrammatic displays for engineered systems: Effects on human performance in interacting with malfunctioning systems. *NASA Technical Report No. 29*, July 29, 1988.

8. Pomerantz, J. R., Pristach, E. A., and Carson, C. E. (1989). Attention and object perception. In B. E. Shepp and S. Ballesteros (Eds.), *Object perception: Structure & process*. Hillsdale, NJ: Lawrence Erlbaum Associates.

9. Sanderson, P. M., Flach, J. M., Buttigieg, M. A. and Casey, E. J. (1989). Object displays do not always support integrated task performance. *Human Factors*, *31*, 183-198.

10. Sternberg, (1966). High-speed scanning in human memory. *Science*, *153*, 652-654.

11. Wickens, C. D. (1984). *Engineering psychology and human performance*. Columbus, OH: Charles E. Merrill.

12. Woods, D. D. & Roth, E. M. (1988). Cognitive systems engineering. In M. Helander (Ed.), *Handbook of human-computer interaction*, North-Holland: Elsevier Science Publishers.

THE USE OF ANALYTICAL MODELS IN HUMAN-COMPUTER INTERFACE DESIGN

Leo Gugerty
Lockheed Engineering and Sciences Company

Research directed by Marianne Rudisill, Manager, Human Computer Interaction Lab, NASA JSC.

INTRODUCTION

Researchers in the human-computer interaction (HCI) field commonly advise interface designers to "know the user." Various approaches are currently used to get information about the user into the hands (and mind) of the designer. One approach is to use design guidelines (e.g., NASA Johnson Space Center, 1988) which can incorporate knowledge of human psychological strengths and weaknesses and make them accessible to designers. However, guidelines give only overview information. They do not help the designer to configure the interface for a specific task and specific users (Gould and Lewis, 1985). Another way to know the user is to conduct usability tests (Gould and Lewis, 1985). This involves building prototype interfaces as early as possible in the design process, observing typical users as they work with the prototype, and fixing any observed problems during the next iteration of the design. While effective in making the designer aware of user needs, usability testing adds a significant amount of time to the design of user interfaces.

Recently, a large number of HCI researchers have investigated another way to know the user - building analytical models of the user, which are often implemented as computer models. These models simulate the cognitive processes and task knowledge of the user in ways that allow a researcher or designer to estimate various aspects of an interface's usability, such as when user errors are likely to occur. This information can lead to design improvements. Analytical models can supplement design guidelines by providing designers rigorous ways of analyzing the information-processing requirements of specific tasks (i.e., task analysis). These models offer the potential of improving early designs and replacing some of the early phases of usability testing, thus reducing the cost of interface design.

This paper describes some of the many analytical models that are currently being developed and evaluates the usefulness of analytical models for human-computer interface design. The paper is intended for researchers who are interested in applying models to design and for interface designers. This is a summary of an extensive literature review paper on the use of analytical models in design that is being conducted at the Johnson Space Center's Human-Computer Interaction Laboratory.

The question of whether analytical models can really help interface designers is currently receiving much attention in the field of human-computer interaction. Advocates of model-based design claim that our knowledge of cognitive psychology is becoming sophisticated enough to allow analytical models of the user to play a useful role in interface design (Kieras, 1988; Butler, Bennett, Polson, and Karat, 1989). Modeling proponents suggest that models could be used during interface design in two important ways:

1. Models can help designers conduct a rigorous task analysis, which in turn may help generate design ideas. A number of analytical models (e.g., the GOMS model, Card, Moran, and Newell, 1983) involve specifying the goals, actions, and information requirements of the user's task. Research suggests that these task

analyses can help designers generate effective design ideas.

2. After interface designs have been generated, models can help evaluate their effectiveness. A human-factors psychologist or engineer could work with a designer to build a computer model of how a user would interact with a new interface. This model could be run with various input conditions to predict how long the user will take to perform tasks using the interface, and likely sources of user errors.

The benefits of analytical models are by no means universally accepted in the HCI community. Many HCI researchers and practitioners have questioned the usefulness of models for interface design. Whiteside and Wixon (1987) claim that current models are only applicable to the specific task and context for which they were developed and cannot be applied to new interfaces. Others (e.g., Curtis, Krasner, and Iscoe, 1988; Rossen, Maas, and Kellogg, 1988) suggest that models may not fit in with the needs of design organizations or with the intuitive thinking and informal planning that designers sometimes use.

This paper will focus on computational, analytical models, such as the GOMS model, rather than less formal, verbal models, because the more exact predictions and task descriptions of computational models may be useful to designers. The literature review paper that is summarized here evaluated a number of models in detail, focusing on the empirical evidence for the validity of the models. Empirical validation is important because without it models will not have the credibility to be accepted by design organizations. This paper will briefly describe two analytical models in order to illustrate important conclusions from the literature review. Following this, the paper will discuss some of the practical requirements for using analytical models in complex design organizations such as NASA.

EMPIRICAL EVALUATION OF ILLUSTRATIVE MODELS

GOMS MODEL

The GOMS model was developed as an engineering model to be used by HCI designers, and it has received much more empirical testing than any other analytical model of HCI tasks. Many of the issues concerning the use of GOMS models in design are relevant to other analytical models as well.

GOMS models are applicable to routine cognitive skills. They are best suited for tasks where users make few errors. More open-ended tasks that involve extensive problem solving and frequent user errors (e.g., troubleshooting) are not good candidates for GOMS modeling.

GOMS stands for goals, operators, methods, and selection rules, the four elements of the model. GOMS models are hierarchical. The assumption is that at the highest level people's behavior on a routine computer task can be described by a hierarchy of goals and subgoals. At the most detailed level, behavior is described by operators, which can be physical (such as typing) or mental (such as comparing two words). Operators that are often used together as a unit are built up into methods. For example, one might have a standard method of deleting text in a text editor. Sometimes more than one method can meet a goal and selection rules are used to choose among them.

GOMS models can help an interface designer get a qualitative understanding of the goal structure and information requirements of a task (i.e., a task analysis). In addition, Kieras and Polson (1985) developed a formal implementation of GOMS models, Cognitive Complexity Theory (CCT), that allows designers to make quantitative statements about users' errors, learning time, and performance time for particular interfaces. In CCT, GOMS models are represented as production systems. In a production system the parts of a GOMS model are represented by a series of if-then rules (production rules)

that can be run as a computer simulation model. A number of quantitative metrics can be derived from a CCT production system that, according to proponents of CCT, can be used to predict users' performance on a task (Kieras, 1988; Olson and Olson, in press). For example, task learning time, task performance time, and the number of user errors can be predicted.

To date, GOMS models have not been used to help design a commercial interface. Most empirical studies of GOMS models have been evaluations of existing interfaces that were designed without using GOMS. For example, Bovair, Kieras, and Polson (in press) evaluated GOMS estimates of task performance time for existing interfaces. Using a text editing task, they found that the number of production-system cycles and of certain complex operators (such as looking at the text manuscript) could match performance time fairly well, explaining about 80% of the variability of users' performance times across editing tasks.

It is important to point out that in studies like this data (such as errors and the time to learn and perform tasks) are collected from users of an interface, and statistical techniques (such as regression) are used to determine whether the GOMS predictions match the data. In these studies, GOMS models are not used to make *a priori* predictions of user performance. Rather, the models' estimates of user performance are statistically compared to the empirical data to see how much of the variability in users' performance data can be explained by the model. Although some researchers suggest that GOMS models can be used to make a priori predictions of user performance (Olson and Olson, in press), this has not been done successfully to date.

In addition to evaluations of existing interfaces, a few studies have looked at how GOMS models can be used to generate ideas for redesigning interfaces. These studies take advantage of the fact that GOMS models provide a detailed task analysis (i.e., a representation of the goals, subgoals, and procedural steps) required to perform a task. Elkerton and Palmiter (1989) used a GOMS model of the knowledge required for Hypercard authoring tasks to design a menu-based Hypercard help system that allowed faster information retrieval and that was liked better than the original help system.

This study is important because it shows that GOMS models can be used for more than post-hoc evaluation of existing designs. In this study, the task analyses provided by GOMS models were used to generate computer-related artifacts (in this case, procedural instructions). In addition, these artifacts were generated fairly directly from the task analyses without extensive interpretation or "judgment calls."

To summarize the empirical evaluation of GOMS models, models developed for a single, existing interface can be used in a *post-hoc*, quantitative fashion to explain performance time, learning time, and number of errors with that interface. No one has yet tested whether GOMS models can make accurate quantitative performance predictions for an interface that is still in design. However, encouraging progress has been made in using the task analyses provided by a GOMS model to help generate effective instructions that can be incorporated in help systems and user manuals.

TULLIS' MODEL

The next model to be described has a much narrower range of application than GOMS models and focuses on general psychological processes rather than task analysis. Perhaps because of these differences, this model, developed by Tullis (1984), is better than GOMS at making *a priori* predictions of user performance. Tullis' model focuses on aspects of a display, such as display density, that affect how well people can find information in the display. It emphasizes general processes, such as perceptual grouping, that affect display perception regardless of the content of the display. The effects of task knowledge on display perception (e.g., effects of user expertise) are not considered. Tullis' model is applicable only to alphanumeric dis-

plays that make no use of color or highlighting. The model has been applied to simple search tasks involving displays for airline and motel reservations and for aerospace and military applications (Tullis, 1984).

Based on a literature review, Tullis hypothesized that five factors would affect the usability of alphanumeric displays: overall density, local density, number and size of the perceptual groups, and layout complexity. He developed operational definitions so that quantitative values could be calculated for each factor, given a display layout as input. Then, he conducted an experiment in which subjects searched for information in displays and rated the usefulness of the displays. Regression analyses showed that the five factors could explain subjects' search times and subjective ratings fairly well.

Tullis implemented his regression model in the Display Analysis Program (Tullis, 1986). This program accepts a display layout as input. It outputs quantitative estimates of overall density, local density, number of perceptual groups, and average group size. It also provides graphical output describing the display density analysis and the perceptual groups. Finally, it predicts average search time and subjective ratings for the display.

Tullis (1984) then used his model to predict search times and subjective ratings for a second experiment, using different subjects and displays than the experiment that was used to develop the regression equations. The predicted search times and subjective ratings matched the actual times and ratings fairly well, with a correlation of about 0.64 (r^2) for each variable. The model correctly predicted the displays with the best search time and rating. Tullis' model was also able to predict search times from three previous studies in the literature ($r^2 > 0.63$ in each study) (Tullis, 1984). However, when Tullis' model was tested on tasks more complex than simple display search, it did not predict subjects' performance well (Schwartz, 1988).

To summarize, Tullis' model is applicable within a limited domain—inexperienced users performing simple search tasks involving alphanumeric displays. Within this domain, however, the model's performance is impressive. Tullis has taken the step that GOMS users have neglected and used his model to predict performance for displays and subjects different from the ones on which the model was developed. The model was able to predict well in these cases. One disadvantage of Tullis' model is that it neglects cognitive factors affecting display perception, such as the effect of a user's task knowledge.

CONCLUSION: EMPIRICAL EVALUATION OF ANALYTICAL MODELS

Earlier in the paper, it was suggested that analytical models could be used in interface design in two ways. The first of these involves using models early in the design process to conduct rigorous task analyses, which are then used to generate ideas for preliminary designs (e.g., menu structures). The second potential use of models occurs later in the design process, after preliminary designs have been developed. In this case models are used to evaluate designs by making quantitative predictions about expected user performance given a particular design.

The empirical evidence considered in the literature review, and summarized here, suggests that, except for one model with a narrow range of application, there is no empirical evidence that analytical models can predict user performance on a new interface. There is some encouraging evidence that analytic models used for task analysis can help in the process of generating designs; however, this conclusion is based on only a few studies. The review of the empirical evidence suggests, then, that future research aimed at demonstrating model-based improvements in interfaces should focus on three areas:

- Replicating and extending the studies of model-based interface redesign (e.g., Elkerton and Palmiter, 1989).
- Demonstrating model-based interface design for a new interface.

- Demonstrating the predictive use of models to evaluate preliminary designs.

Based on the empirical evidence to date, the first two of these would be the most promising avenues of research.

What are some possible reasons for the failure of models to accurately predict performance with a new interface? It may be that critics such as Whiteside and Wixon (1987) are correct in that people's procedures, goals, and cognitive operators are too context specific to allow prediction in a context as different as a new interface. A large body of research in cognitive psychology suggests that experts' performance in a particular domain is largely dependent on domain-specific knowledge, as opposed to general-purpose cognitive skills (Chi, Glaser, and Rees, 1982; Glaser, 1984). And models such as GOMS focus primarily on the task-specific knowledge of experienced users. It is interesting that the model that was able to predict user performance on a slightly different interface (Tullis') is not a task analytic model. Tullis' model focuses on general perceptual abilities. This suggests that in order to predict performance for new interfaces, task analytic models must include more explicit representation of how general purpose cognitive characteristics (such as working memory limitations) affect user performance.

An addition should be made to the above list of research areas. This suggestion is based on the fact that there are no empirically validated models that can describe HCI tasks involving higher-level cognitive processes such as problem solving. However, space-related computer systems are rapidly becoming intelligent enough to assist people in complex tasks, such as medical diagnosis and scientific research, which involve more complex cognition. Models are currently being developed with the goal of describing these more complex tasks in a way that is useful to interface designers. An example is the Programmable User Models (PUMs) (Young and Whittington, 1990). However, most of these models have not been empirically validated.

A fourth area of further research, then, is:

- Developing and testing models of complex HCI tasks involving high-level cognitive processes.

USING MODELS IN DESIGN ORGANIZATIONS

So far, this paper has focused on whether analytical models can improve interface designs. However, even if models were conclusively demonstrated to improve interfaces, this would still not ensure their use by design organizations such as NASA. What is needed is evidence for the usefulness as well as the validity of models. That is, it must be shown that models can meet the needs of individual designers (e.g., preferred design methods) and of design organizations (e.g., cost, scheduling, and personnel constraints).

With respect to individual designers, an understanding of the various ways that designers generate, develop, and evaluate ideas is needed. Analytical models would be provided to designers as detailed procedures or as software tools. The principle of considering the cognitive and motivational processes of users applies to model developers just as it does to the designers of other software tools. In short, designers are users too. Therefore, if model developers want their models to be used in actual design projects, they must either construct their models to fit in with the preferred design processes of designers or provide ways of training designers to use the models.

But decisions regarding the commercial use of models are made by managers, not by individual designers. Therefore, models also must be shown to meet the multifaceted needs of design organizations, for example, cost, schedule, and personnel requirements. This section will discuss the problems that must be overcome before analytical models are accepted by designers and their work organizations.

NEEDS OF INDIVIDUAL DESIGNERS

Two studies conducted by Curtis and his colleagues showed that major difficulties in software design are caused by a lack of application-domain knowledge on the part of designers. (Curtis et al., 1988; Guindon, Krasner, and Curtis, 1987). The analogous problem in the case of interface design would be a lack of knowledge of the user's task. When Rosson et al., (1988) interviewed interface designers about the techniques they used to generate design ideas, they found that the most frequently mentioned techniques (about 30%) were for analyzing the user's task. Most of this task analysis involved informal techniques, such as interviewing users or generating a task scenario.

These findings present both an opportunity and an obstacle to the use of models by interface designers. First, since designers often lack knowledge of the user's task and spend a large amount of effort getting it, they might see the usefulness of task analytic models such as GOMS. The potential obstacle is that designers may prefer to stick with their informal techniques, instead of the more rigorous task analytic models. Rosson et al., suggest that tools to aid in idea generation should primarily support designers' informal techniques. Lewis, Polson, Wharton, and Rieman (1990) offer an interesting way of combining formal modeling with a technique currently used by software designers—design walkthroughs. They developed a formal model of initial learning and problem solving in HCI tasks, and then derived from the model a set of structured questions (a cognitive walkthrough) that can be used to evaluate the usability of an interface.

This discussion presents only an example of the kind of issues that need to be considered regarding the needs of individual designers. Further research is needed on the cognitive and motivational processes of designers and what these processes suggest about the design of analytic models.

NEEDS OF DESIGN ORGANIZATIONS

The Curtis et al., (1988) study mentioned above also considered the organizational aspects of software design. In addition, Grudin and Poltrock (1989) conducted an extensive interview study of the organizational factors affecting interface design. Some of the findings of these studies that relate to the use of analytical models are discussed below.

An important characteristic of many computer-system design organizations is complexity. Many groups may contribute to a final design product: interface and system designers, human factors personnel, training developers, technical writers, and users (e.g., astronauts). Curtis et al., (1988) noted a wide variety of communications problems that resulted because of this organizational complexity. One such problem arises when groups interpret shared information differently because of differences in background knowledge. This could easily cause problems, for example, if the people in an organization who are experienced with modeling (e.g., a designer or human factors expert) have to communicate the results of a modeling analysis to a project manager. A possible solution to this problem of misinterpretation is for model developers to make the structure and outputs of their models as clear as possible.

In addition to communication problems, another problem arising from the variety of roles in design organizations has to do with personnel and training. A manager considering the use of models on a design project faces a number of questions along these lines. Can existing personnel do the modeling (e.g., designers or human factors personnel)? How much training will they require? If new personnel must be hired, what kinds of background must they have? Model developers must have answers to these questions.

One answer comes from the work of Kieras (1988). He has developed and published a procedure for building GOMS models. Informal testing showed that computer science

undergraduates could use this procedure to generate GOMS models and make usability predictions "with reasonable facility." More than this is necessary, however. Validation studies must be done to test whether the personnel that would use models in design organizations can build models that make the same kinds of predictions as the experts who initially developed the model. These studies should also document the kind of training necessary to achieve these ends.

In addition to complexity, other characteristics of design organizations that affect their openness to modeling are strict project scheduling and a concern with monetary costs. Detailed estimates are needed of the time and money costs of using analytical models in commercial design.

CONCLUSION: THE USE OF ANALYTICAL MODELS IN INTERFACE DESIGN

Can the use of analytical models be recommended to interface designers? Based on the empirical research summarized here, the answer is: not at this time. There are too many unanswered questions concerning the validity of models and their ability to meet the practical needs of design organizations. However, some of the research described here suggests that models can be of practical use to designers in the near future. Of special interest is the research that used models as task analytic tools to generate interface design ideas (e.g., Elkerton and Palmiter, 1989).

This paper has suggested research and development that is necessary in order for analytical models to be accepted by complex design organizations. These suggestions are summarized in Table 1. It seems that the empirical research on analytical models gives good reason to pursue the research and development goals outlined here.

TABLE 1.

Methods of Increasing the Use of Analytical Models in Interface Design

Demonstrate design improvements:
- Validate model-based interface redesign.
- Validate model-based interface design.
- Validate predictive use of models to evaluate preliminary designs.
- Develop and validate models of complex HCI tasks involving high-level cognitive processes.

Meet the needs of individual designers:
- Study the design methods and cognitive processes of individual designers.
- Change the models and/or develop training materials to ensure that models fit in with designers' methods and cognitive processes.

Meet the needs of design organizations:
- Make models' structure and outputs easily interpretable.
- Develop means of training designers to use models. Validate that this training works and document the costs of training.
- Document the time and monetary costs of using models.

ANALYTICAL MODELS AND SPACE-RELATED INTERFACE DESIGN

So far, this paper has provided a general analysis of the use of analytical models in human-computer interface design. How much of this analysis is applicable to the design of space-related interfaces? The Human-Computer Interaction Laboratory (HCIL) at the Johnson Space Center is currently conducting preliminary task analyses for the tasks required on a long-duration space mission, such as a mission to Mars (Gugerty and Murthy, in preparation). This work suggests that the range of tasks on such a mission is quite broad—ranging from reading to controlling complex equipment to conducting scientific research. The possible information technologies for long-term missions are also quite diverse, for example, workstations for supervisory control, graphics workstations for scientific research, computer-supported group meetings, medical expert systems, and virtual workstations for

telerobotic control. It seems that space-related tasks are diverse enough to span almost the entire range of human-computer interaction tasks. Therefore, the general analysis of this paper will be applicable to space-related tasks in most cases.

One project in the JSC HCIL is focusing on the use of analytical models in designing medical decision support systems for space crews. This project is following up on the work of Elkerton and Palmiter (1989) in which GOMS was used as a task analytic model to help generate interface design ideas. One medical task that space crew members will face is learning or relearning medical procedures from computer displays. This project will test whether building GOMS models of medical procedures can help interface designers build better interfaces for displaying this procedural information. The GOMS approach will be compared with other methods of task analysis, including psychological scaling techniques such as the Pathfinder algorithm (McDonald and Schvaneveldt, 1988).

REFERENCES

1. Bovair, S., Kieras, D. E., and Polson, P. G. The acquisition and performance of text editing skill: A production system analysis, *Human Computer Interaction*, in press.

2. Butler, K, Bennett, J., Polson, P., and Karat, J. Report on the workshop on analytical models, *SIGCHI Bulletin*, 20, 4, 1989, pp. 63-79.

3. Card, S. K., Moran, T., and Newell, A. *The psychology of human-computer interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1983.

4. Chi, M. T. H., Glaser, R., and Rees, E. Expertise in problem solving, in R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1), Lawrence Erlbaum Associates, Hillsdale, NJ, 1982.

5. Curtis, B., Krasner, H., and Iscoe, N. A field study of the software design process for large systems, *Communications of the ACM,* 31,11, 1988, pp. 1268-1287.

6. Elkerton, J. and Palmiter, S., Designing help systems using the GOMS model: An information retrieval evaluation, *Proceedings of the Human-Factors Society 33rd Annual Meeting*, Santa Monica, CA, 1989.

7. Glaser, R. Education and thinking: The role of knowledge, *American Psychologist*, 39, 2, 1984, pp. 93-104.

8. Gould, J. D. and Lewis, C. H. Designing for usability and what designers think, *Communications of the ACM*, 28, 3, 1985, pp. 300-311.

9. Grudin, J. and Poltrock, S. E. User interface design in large corporations: Coordination and communication across disciplines. In K. Bice and C Lewis (Eds.), *Proceedings of the CHI 89 Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY 1989.

10. Gugerty, L. J. and Murthy, P. (in preparation). Using task analysis to guide information-technology development for future space missions, Paper to be presented at the AIAA Space Programs and Technologies Conference, September, 1990.

11. Guindon, R., Krasner, H., and Curtis, B. Breakdowns and processes during the early activities of software design by professionals. In G. M. Olson, S. Sheppard, and E. Soloway (Eds.), *Empirical studies of programmers: Second workshop*, Ablex Publishing Corp., Norwood, NJ, 1987.

12. Kieras, D. E. Towards a practical GOMS model methodology for user

interface design. In M. Helander (Ed.), *Handbook of human-computer interaction,* Elsevier Science Publishers, Amsterdam, 1988.

13. Kieras, D. E. and Polson, P G. An approach to the formal analysis of user complexity, *International Journal of Man-Machine Studies,* 22, 1985, pp. 365-394.

14. Lewis, C., Polson, P., Wharton, C., and Rieman, J. Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces, in J. C. Chew and J. Whiteside (Eds.), *Proceedings of the CHI 90 Conference on Human Factors in Computing. Systems*, Association for Computing Machinery, New York, NY, 1990.

15. McDonald, J. E. and Schvaneveldt, R. W. The application of user knowledge to interface design. In R. Guindon (Ed.), *Cognitive science and its applications for human-computer interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1988.

16. NASA Johnson Space Center, Space Station Freedom Program human-computer interface guide (USE 1000, Version 2.1), Johnson Space Center, Houston, TX, 1988.

17. Olson, J. R. and Olson, G. M. The growth of cognitive modeling in human-computer interaction since GOMS, *Human-Computer Interaction*, in press.

18. Poltrock, S. E. Innovation in user interface development: Obstacles and opportunities. In K. Bice and C. Lewis (Eds.), *Proceedings of the CHI 89 Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, 1989.

19. Rosson, M. B., Maass, S., and Kellogg, W. A. The designer as user: Building requirements for design tools from design practice, *Communications of the ACM,* 31, 11, 1988, pp. 1288-1299.

20. Schwartz, D. R. The impact of task characteristics on display format effects, *Proceedings of the Human Factors Society - 32nd Annual Meeting*, Human Factors Society, Santa Monica, CA, 1988.

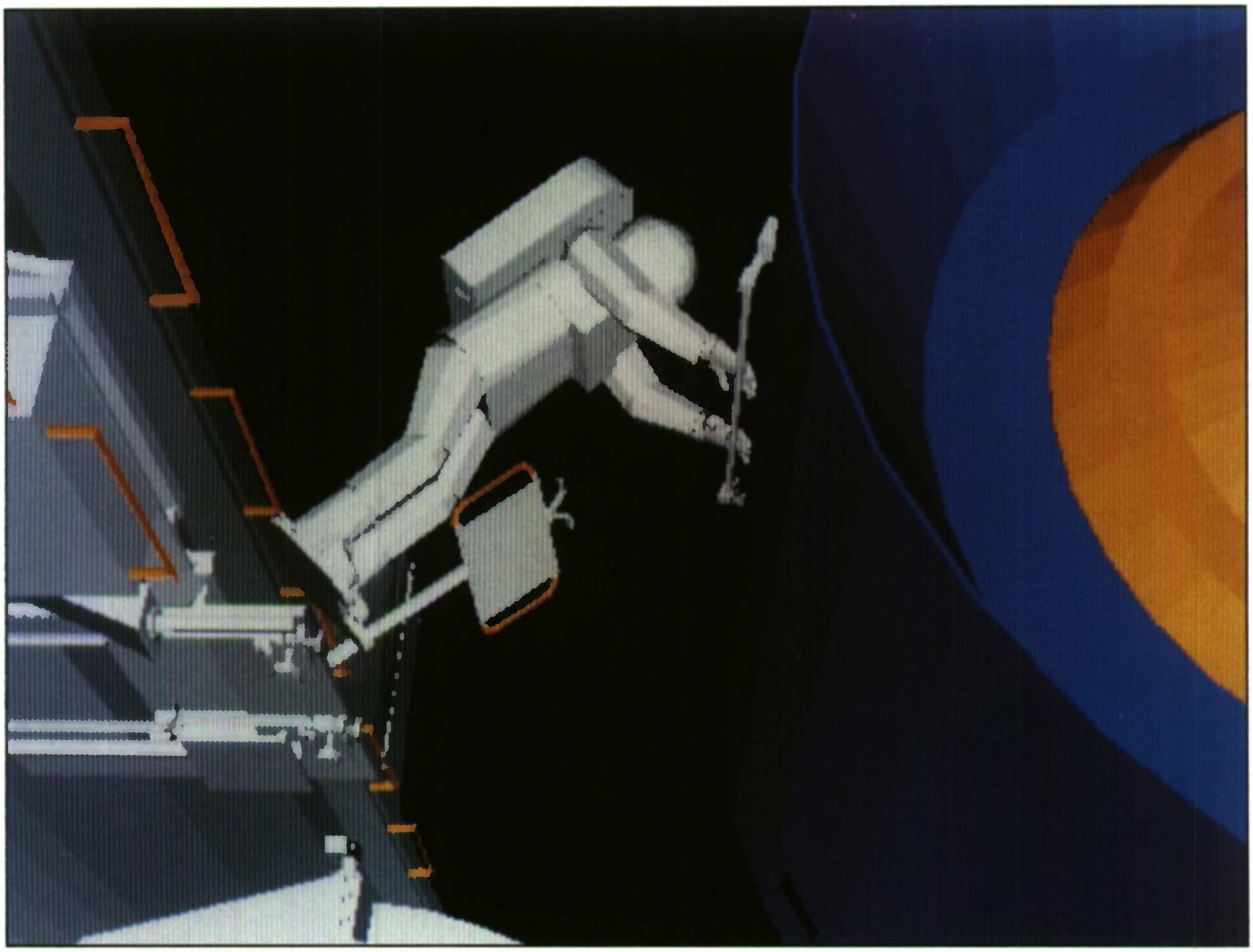
21. Tullis, T. S. Predicting the usability of alphanumeric displays, Doctoral Dissertation, Rice University, The Report Store, Lawrence, Kansas, 1984.

22. Tullis, T. S. A system for evaluating screen formats, *Proceedings of the Human-Factors Society 30th Annual Meeting*, Human Factors Society, Santa Monica, CA, 1986.

23. Whiteside, J. and Wixon, D. Discussion: Improving human-computer interaction - A quest for cognitive science. In J. M. Carroll (Ed.), *Interfacing thought: Cognitive aspects of human-computer interaction*, MIT Press, Cambridge, MA, 1987.

24. Young, R. M. and Whittington, J. Using a knowledge analysis to predict conceptual errors in text-editor usage, in J. C. Chew and J. Whiteside (Eds.), *Proceedings of the CHI 90 Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, 1990.

Space Habitability



A three-dimensional interactive computer graphics package called PLAID is used to address human factors issues in spacecraft design and mission planning. Premission studies produced this PLAID rendition to show where an EVA astronaut would stand while restraining a satellite manually and what the IVA crewmember would be able to see from the window.

(See cover for the actual photo taken during mission from aft crew station.)

USING COMPUTER GRAPHICS TO DESIGN SPACE STATION *FREEDOM* VIEWING

Betty S. Goldsberry, Buddy O. Lippert, and Sandra D. McKee
Lockheed Engineering and Sciences Company

James L. Lewis, Jr. and Francis E. Mount
NASA Lyndon B. Johnson Space Center

INTRODUCTION

Viewing requirements were identified early in the Space Station Freedom program for both direct viewing via windows and indirect viewing via cameras and closed-circuit television (CCTV). These requirements reside in NASA Program Definition and Requirements Document (PDRD), Section 3: Space Station Systems Requirements.

Currently, analyses are addressing the feasibility of direct and indirect viewing. The goal of these analyses is to determine the optimum locations for the windows, cameras, and CCTVs in order to meet established requirements, to adequately support space station assembly, and to operate on-board equipment.

PLAID, a three-dimensional computer graphics program developed at NASA JSC, was selected for use as the major tool in these analyses. PLAID provides the capability to simulate the assembly of the station, as well as to examine operations as the station evolves. This program has been used successfully as a tool to analyze general viewing conditions for many Space Shuttle elements and can be used for virtually all Space Station components. Additionally, PLAID provides the ability to integrate an anthropometric scale-modeled human (representing a crewmember) with interior and exterior architecture.

BACKGROUND

COMPUTER SIMULATION

The design of a computer simulation system, such as PLAID, that includes human models is a complex process. Total system performance is dependent on the accuracy of the models generated as well as the interactions between humans, hardware, and software. Model requirements can be based on workspace geometry, figure anthropometry, strength/force characteristics, and reach envelopes. If results of simulation analyses are to be valuable, development of the models must be based on a valid representation of the environment. Also, procedures must be employed that make available alternative human/system designs. Ultimately, quantitative predictions of events and behaviors in response to realistic operating conditions for various design alternatives must be compared in order to select the optimum design characteristics.

While there are many ways of quantitatively predicting human behavior and performance, computer simulation of the human operator in a mission context represents a method that is internally consistent and compatible with other contemporary system engineering evaluative techniques. Computer simulations usually resolve many design and development problems in an effective manner earlier than experiments with human subjects. Even though simulation does not eliminate the need for empirical tests, properly exploited, it can help to focus test time and energy on appropriate issues and potential problems.

The analyst interested in evaluating human performance can use computer simulation to construct a graphic human model in much the same way that an airplane or automobile designer can model a vehicle. Then, the model can be used to test various design concepts relevant to human-system interactions and integrations. Actually, a computer simulation

system can be developed long before a prototype or test facility can be made operational.

PLAID MAN-MODELING SYSTEM

PLAID uses anthropometric data collected from astronaut candidates in the JSC Anthropometrics and Biomechanics Laboratory to generate human models with realistic joint limits and user-specified size characteristics. A high-fidelity model of the Space Shuttle Remote Manipulator System (RMS) is also contained in the PLAID database. For many years, the models have facilitated the comparison of reach envelopes for humans and remote manipulators in order to determine which space operations are feasible for astronaut extravehicular activities (EVA) and which are more appropriately accomplished by robotics applications. PLAID has also played a major role in satellite retrieval execution, equipment failure diagnosis, and vehicle damage prediction.

PLAID is a three-dimensional-solids modeling system which generates computer models that can be examined from an infinite number of viewpoints. This feature allows the system user to position its viewpoint where the eyes of the crewmember or lens of the camera would be located. Then, subsequent analyses can reveal what a crewmember can or cannot see. The PLAID system has taken into consideration that the visual environment of space is different from a normal Earth atmosphere and has incorporated appropriate contrast ratios, shadowing, and light scattering for the space environment.

SPACE STATION

PROCEDURE

Since viewing tasks vary in complexity, relevant viewing requirements and proposed window locations were identified. Using the PLAID system, geometric models of Space Station Freedom elements and configuration models were created. These models reflected up-to-date Space Station Freedom architectural information for internal and external elements. Also, proposed window and camera locations were integrated with the models. Then, computer graphics of various viewing scenarios were generated which revealed fields-of-view for selected locations of interest. Finally, the graphics generated by the system were analyzed relative to published design requirements and specifications.

The simulation is influenced by more than a three-dimensional layout of the environment. In particular, task sequence and task time are important ingredients. For example, if the humans must monitor exterior facilities while operating interior controls, their fields-of-view will be decreased because their positions will probably be farther from the window. Shadowing can also reduce visibility.

The simulation helped analysts evaluate the proposed window locations and identify problem areas. Recommended window placements were determined by the integration of the view available from a particular position with various other factors such as operations requirements, anthropometric clearances, traffic flow, and window accessibility.

DATA

Data in the form of PLAID graphics were generated and maintained at the JSC/Man-Systems Division. Figures 1 through 4, illustrated on the following page, exemplify the types of data used in the analyses. Figure 1 was generated to determine the viewing clearance needed if an airlock were placed near the window.

Figure 2 illustrates that the cupola will accommodate two crewmembers. Note the design requirement that the Crew Escape Rescue Vehicle (CERV) be visible from the cupola is met. Displays that share the crewmembers' fields-of-view, are shown.

Figures 3 and 4 illustrate that viewing of the external environment can be accomplished. Analyses have determined that Earth viewing or optic experiments could be obstructed by truss work from some locations.

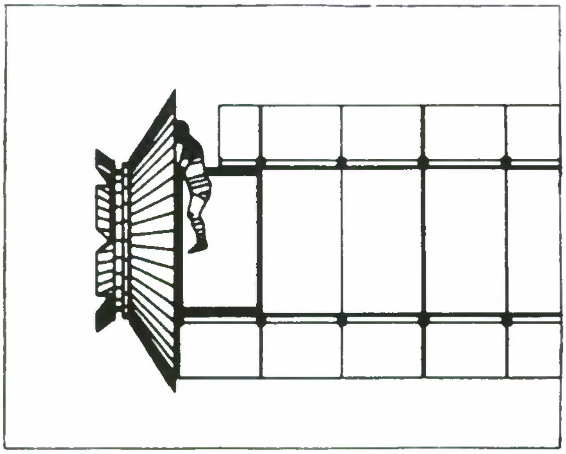


Figure 1. A 95th percentile male crewmember looking out of the cone-shaped end portion of a module (endcone) window.

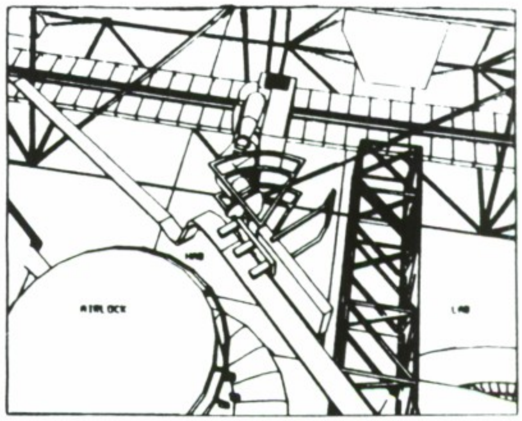


Figure 3. A view of extravehicular activity (EVA).

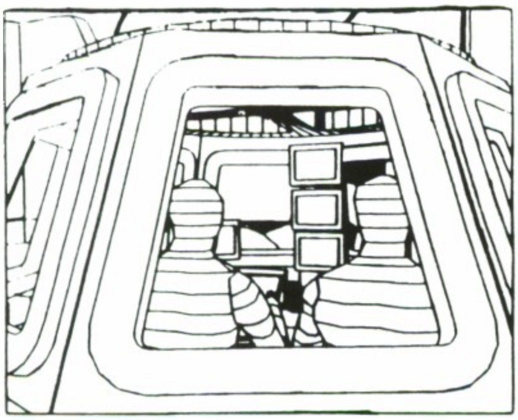


Figure 2. Two 50th percentile males in a cupola.

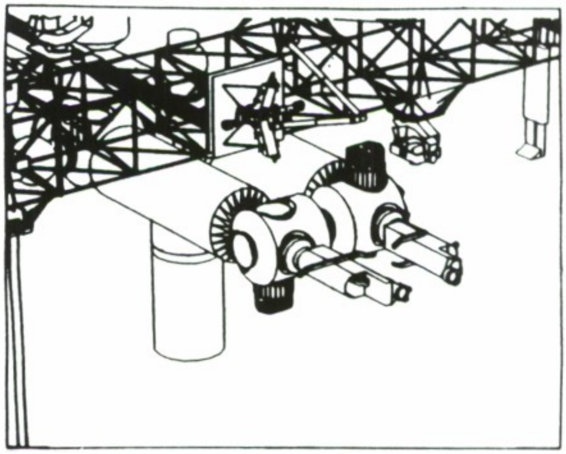


Figure 4. An Orbiter window or external camera view during approach or departure.

DOCUMENTATION

Comprehensive documentation of the Space Station Viewing Analyses results will be organized in a series of volumes. Volume 1 covers viewing of the Space Station *Freedom* assembly sequence as seen through the Orbiter windows and CCTVs. Volume 2 uses the Man-Systems candidate topologies to evaluate possible window locations in the U. S. Habitation module. Volume 3 evaluates the candidate window locations as proposed by the European Space Agency (ESA) for the Columbus Module, as well as selected alternate locations. Volume 4 addresses viewing conditions of the National Space Development Agency (NASDA) Japanese Experiment Module (JEM). Volume 5 is dedicated to the discussion of viewing from the nodes and the two cupolas baselined on nodes three and four. Volume 6 addresses indirect viewing and the placement of cameras throughout the station. Volume 7 pertains to the viewing conditions of the U. S. Laboratory module. Volume 8 will

address Phase II Space Station Freedom viewing impacts. The planned order of preparation for these basic volumes is 3, 2, 4, 7, 5, 1, 6, and 8. The first four volumes in this series have been released. Published recommendations from the completed volumes follow.

RECOMMENDATIONS

U. S. Habitation Module

1. A requirement for windows in the crew quarters should be established by the program.

2. Two viewing stations, currently proposed, should be relocated to positions that will provide unobstructed views.

3. Modular interior layout and wardroom area window placement should be integrated such that recreational viewing (i.e., viewing of the Earth limb) is optimized. Selected port rack locations are well suited for wardroom area windows.

4. Windows should be placed in the portion of the quadrant where the eyes are most likely to be, while maintaining compliance with the local vertical.

U. S. Laboratory Module

1. Evaluate scientific celestial viewing from the zenith (-Z) oriented windows in the module to insure minimal interference of Space Station elements.

2. A proposed floor window is suitable for Earth viewing.

3. Optimum high-oblique viewing in the -Y direction for scientific uses can be achieved from a starboard window. Exact placement of the window below the module centerline should be determined based upon intended use of the window.

ESA Columbus Module (CM)

1. Design of the module should provide ample space for movement, interactions, interfaces, and other human factors considerations.

2. Orient a minimum of two windows in the CM in the +X and -X directions.

3. Rotate the aft endcone at least 45 degrees from 12 o'clock position to meet Space Station Program 30000 requirements and support recreational viewing.

4. Rotate the window location in the forward endcone of the CM 45 degrees clockwise.

5. Consider the need for celestial viewing from the CM in the overall station configuration.

6. Establish commonality of window design characteristics in all modules.

7. Supplement the view through the hatch viewport by providing interior video which could also be used to remotely monitor activities within the modules.

NASDA Japanese Experiment Module

1. Supplement direct viewing of Mobile Servicing Center (MSC) operations by using the CCTV for indirect viewing.

2. Supplement direct viewing of the Experiment Logistics Module (ELM) transfer and berthing/deberthing operation by using the CCTV for indirect viewing and for supporting direct viewing from the hatch viewport beneath the JEM berthing ring during the berthing process.

3. Retain proposed windows in the aft bulkhead.

4. Retain baseline with no windows in the forward bulkhead, since no meaningful view could be achieved.

5. Retain floor window for Earth observation.

6. For purposes of celestial viewing, use the viewport between the JEM and ELM during absence of the ELM. Further information is

needed concerning (a) duration of the ELM's absence and (b) external covering of the viewpoint during ELM's absence.

CONCLUSION

VIEWING ANALYSIS

PLAID allows for an in-depth analysis of the impact of direct and indirect viewing on architectural design and human performance issues. Specifically, window and video location selections have been facilitated by the use of this computer simulation system. This system permits the integration of selected window/camera locations with the internal architecture, the validation of crew activity sequences, and the determination of necessary body movements required to accomplish viewing tasks. Through findings in these analyses, analysts have been able to write a Change Request (CR) affecting high level Space Station Program documents. Requests address placement issues in the modules. Future volumes will address indirect viewing, cupola viewing, node windows, and Space Station assembly.

PLAID SYSTEM

Development of a more interactive, real-time system which uses natural language will enhance capabilities of the system. Expansion of the human-modeling capability will include a population range from the 95th percentile to the 5th percentile female. Currently, work in this area is being conducted for JSC and other users at the University of Pennsylvania under the direction of Norman Badler, PhD.

Also, the capability of generating video animation sequences is being added to PLAID. This feature will allow for simulation of a complete task sequence rather than a "snapshot" approach. Animation will provide for a more comprehensive evaluation of crewmember activities, as well as reveal the relationship between a series of contiguous activities. By following the action flow, the system user can isolate spatial interferences, procedural inconsistencies, and crewmember interactions.

Finally, a dynamic program extension is planned that will allow system users to more accurately evaluate factors such as force and path trajectory. In addition, an efficient ray trace algorithm would enhance lighting study capabilities.

ACKNOWLEDGEMENTS

These viewing analyses could not have been accomplished without the NASA Graphics Analysis Facility, Facility Manager Linda Orr, and Lockheed technical support.

REFERENCES

1. Mount, Frances E. and McKee, Sandra. Space Station Freedom Viewing Analysis: Volumes 2, 3, 4, and 7, JSC No. 32089, NASA Johnson Space Center, 1989.

2. Mount, Frances E. and Lewis, James L. Space Station Viewing Requirements (1986). 1986 SAE Aerospace Technology Conference and Exposition.

A SIMULATION SYSTEM FOR SPACE STATION EXTRAVEHICULAR ACTIVITY

Jose A. Marmolejo
NASA Lyndon B. Johnson Space Center

Chip Shepherd
Lockheed Engineering and Sciences Company

Research directed by Barbara Woolford, Flight Crew Support Division, NASA JSC.

INTRODUCTION

America's next major step into space will be the construction of a permanently manned Space Station which is currently under development and scheduled for full operation in the mid-1990s. Most of the construction of the Space Station will be performed over several flights by suited crewmembers during an extravehicular activity (EVA) from the Space Shuttle. Once fully operational, EVAs will be performed from the Space Station on a routine basis to provide, among other services, maintenance and repair operations of satellites currently in Earth orbit.

BACKGROUND

When a crewmember ventures outside the spacecraft, an extravehicular activity (EVA) is performed. To perform an EVA from the Space Shuttle, an astronaut must don an extravehicular mobility unit (EMU) — a space suit assembly and portable life-support backpack that provides, respectively, the pressure retention and habitable atmosphere that a human requires to perform a productive umbilical-free EVA in the vacuum of space. Attached to the front of the Shuttle EMU shown in Figure 1 is a display and control module (DCM), a large chestpack with life-support controls and a 12-character red LED display located just beneath the helmet visor. Through the DCM display, the astronaut is able to read various life-support parameters (e.g., oxygen pressure, suit pressure, etc.). These parameters are continuously measured and monitored by a caution and warning system (CWS) microcomputer located within the backpack. If any anomalous conditions are detected, caution or warning messages are generated and then relayed from the CWS to the EVA crewmember via the DCM display. Should any of these messages be reported by the CWS, the astronaut then refers to corrective instructions in the EMU cuff checklist, a "flip-through" reference booklet attached to the wrist of the suit. (In addition to these corrective instructions, the cuff checklist also contains information about general mission procedures and EVA equipment operation.)

LIMITATIONS TO THE SHUTTLE EMU INFORMATION SYSTEM

Although several spectacular EVAs have been performed from the Shuttle, a number of limitations with the Shuttle EMU information-exchange system have been identified. These limitations relate to the DCM visibility, information accessibility, information capacity, and the DCM size. Both the display and the controls on the DCM are difficult to view. The location of the DCM display requires the crewmember to look down at an uncomfortable angle. Since its viewing distance is so small, some astronauts require the use of a special lens to read the characters. In addition, bright environments "wash out" the red LED characters, sometimes forcing the crewmember to cup his hands over the display for viewing. Difficulties also exist in viewing the controls located on the front of the DCM, requiring the EVA crewmember to wear a wrist mirror to read the reflected images.

Information retrieval during EVA must be manually sequenced. The astronaut must

toggle a switch to receive each 12-character line of DCM data, and referring to the cuff checklist is a two-hand operation.

The amount of data that may be presented is severely limited. The entire library available during EVA consists of a number of twelve-character messages (an astronaut would have to toggle 24 times just to read this paragraph) and 40 to 50 pages of a 3.25-inch x 4.5 inch cuff checklist.

The DCM not only invades the astronaut's prime work envelope (i. e., the area directly in front of the chest), but its mere presence has restricted improvements to the reach capability and arm mobility of the EMU.

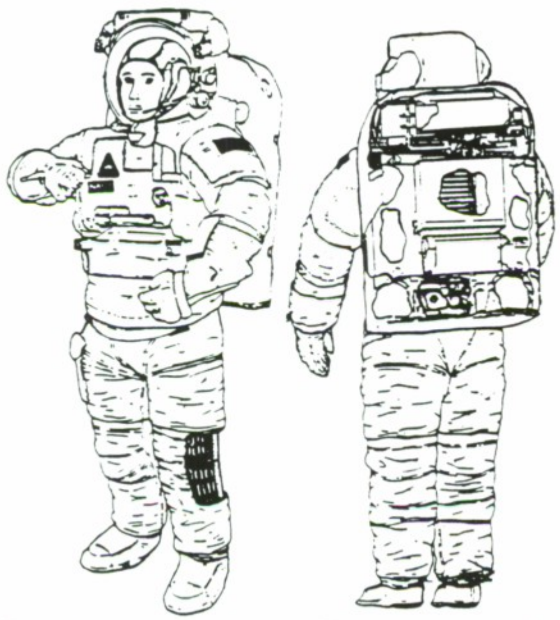


Figure 1. Space Shuttle Extravehicular Mobility Unit (EMU).

INFORMATION SYSTEM IMPROVEMENTS

Although the DCM and cuff checklist are suitable for the occasional EVAs performed from the Shuttle, the nearly routine EVA anticipated for the Space Station dictates that productivity be maximized. Ideally for Space Station EVA, the access to information should be a "hands-free" operation, especially for labor-intensive EVA scenarios such as satellite servicing, unplanned maintenance, or emergency operations. Furthermore, the display medium should support text, graphical, and video formats in all EVA environments. To meet these ambitious goals, NASA is developing a voice recognition and control system to provide "hands-free" operations and a helmet-mounted projection device to display information to the EVA crewmember.

Although space is inherently a harsh environment, the rather benign conditions inside the EMU facilitate the use of speech recognition technology. First, sounds within the EMU are absent of noises that commonly cause problems for speech recognizers, such as background voices and extraneous clamor found at industrial workstations or breathing noises inherent with masked fighter pilots. The EMU helmet contains only fan and airflow noises which are high-frequency sounds that can be easily filtered. Second, an extensive vocabulary is not required for the EMU since a vocabulary of about 50 words will suffice for the procedurally oriented tasks of EVA. Furthermore, even though a connected-word recognizer may allow more flexibility, an isolated-word recognizer is all that is necessary. Finally, the voice recognition system need not be speaker-independent to allow multiple users since each EVA-qualified crewmember aboard the Space Station will have an EMU. Similar to military helmet-mounted and heads-up displays, the EMU helmet-mounted display (HMD) consists of projection optics and image/illumination source electronics that permit convenient viewing of text, graphics, and video on a transparent display deposited onto the helmet visor. With a transparent screen, the astronaut can read information overlaid onto the real world when the HMD is active. When the HMD is off, the screen is practically invisible. Displayed images are located conveniently above the astronaut's prime field-of-view to prevent visual interference while performing EVA tasks.

Through control system electronics, the voice recognizer can select imagery on the HMD, thus providing a powerful and valuable tool for the Station EVA astronaut (other functions are also available). The combined system will

replace the need for the DCM, the cuff checklist, and the EMU wrist mirror, thus enhancing EVA productivity while sacrificing none of their services.

With the HMD and the aid of a voice recognition system, almost unlimited amounts of information are conveniently available to the astronaut during the EVA. The astronaut may review EMU status information, receive EMU alarms with corrective action, obtain general EVA operational information (previously supplied by the cuff checklist), control EMU switches by voice, monitor transmissions from the astronaut's personal TV camera, and access the Station's main computer (which can provide data retrieval from onboard memory or a ground database).

NASA EVA SIMULATION

At the NASA Johnson Space Center, the data available to a Space Station EVA astronaut may be simulated in an Advanced EVA Systems Laboratory. This simulation concentrates on the anticipated Space Station EMU data logic flow required to efficiently provide the EVA crewmember with various informational types while maximizing the usefulness of the voice recognition/control system and helmet mounted display. The simulation program is run on a MacIntosh II and operates in conjunction with a bench model of the HMD and a speaker-dependent voice recognition system. The simulation allows the user to perform the pre-EVA series of checklists required to exit the spacecraft, offers the user all informational data types (including alarms) available during the EVA, and provides the series of checklists to return safely inside the vehicle. Below is a brief description of the simulation program.

PRE-EVA

Pre-EVA procedures are performed within the airlock prior to exiting the spacecraft. After the EMU is donned and powered up, the HMD supports a preparation mode where the astronaut user is led through a series of checks prior to depressurizing the airlock. Detailed instructions including illustrations for depressurization and subsequent airlock egress are provided. In the simulation, the pre-EVA HMD screens are split into three fields (excluding the header). The top field contains procedural information, the middle field contains a series of switches that may be manipulated by voice, and the bottom field lists valid functions selectable by voice. The pre-EVA checks and switch manipulations are controlled by voice commands. The following is the sequence of events required to perform the pre-EVA operation. (Note the quoted commands spoken by the user.)

(1) Simulating power applied to the EMU, the HMD will display a power restart message and a Systems Check of the EMU (See Figure 2).

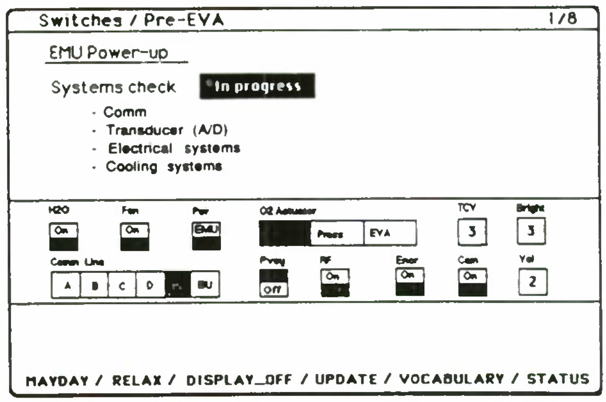


Figure 2. EMU Power-up and Systems Check

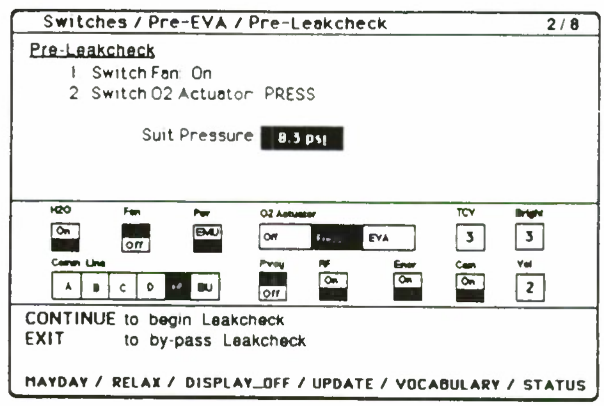


Figure 3. Pre-EVA Checklist

(2) Following the systems status check (and if OK), the first page of the pre-EVA file is shown (see Figure 3). The crewmember is directed to turn the EMU fan on and then to

pressurize the suit. To activate the fan, the crewmember simply says "FAN ON." The FAN block in the middle field will then toggle position. In a similar fashion, the suit is pressurized by first selecting "O2 ACTUATOR" followed by "PRESS." After the suit is fully pressurized, the next page, suit leak check, will be shown.

(3) A leak check (to determine suit integrity) is performed automatically in the simulation. During the leak check simulation, a status of the check including time left, suit pressure, and O2 actuator positions are displayed (See Figure 4). The leak check may be aborted in progress by saying "EXIT." An aborted leak check may be either restarted ("ACKNOWLEDGE") or by-passed entirely ("CONTINUE"). An unaborted leak check with positive results is shown in Figure 5. (Note that a second test, if desired, can be conducted by saying "ACKNOWLEDGE.")

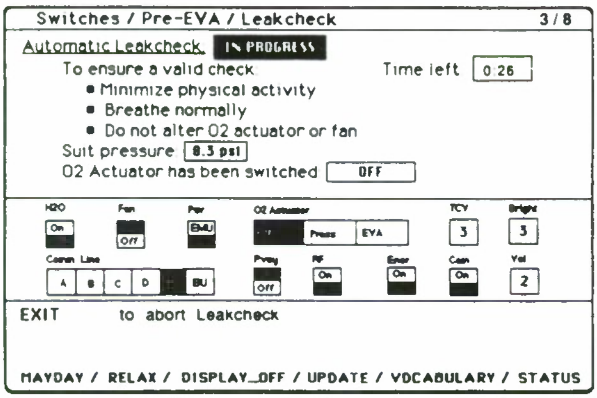


Figure 4. Pre-EVA Leak Check (in-progress)

(4) By saying "CONTINUE," the crew-member will proceed to the Airlock Depress Checklist, a series of procedures to depressurize the airlock to vacuum (see Figure 6). Note the use of graphics. (A "GO BACK" command at this point will display Figure 5, but will *not* redo the actual check unless the previous leak check was aborted or was not satisfactory.) By saying "CONTINUE" again, the crewmember will enter the second page of the checklist (see Figure 7). Note that airlock pressure is located within the checklist step requiring the monitoring.

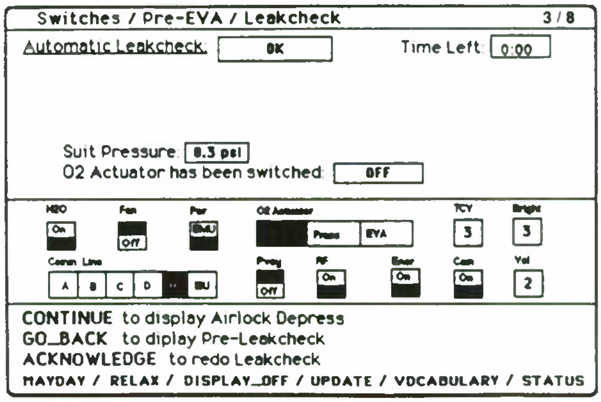


Figure 5. Pre-EVA Leak Check (completed)

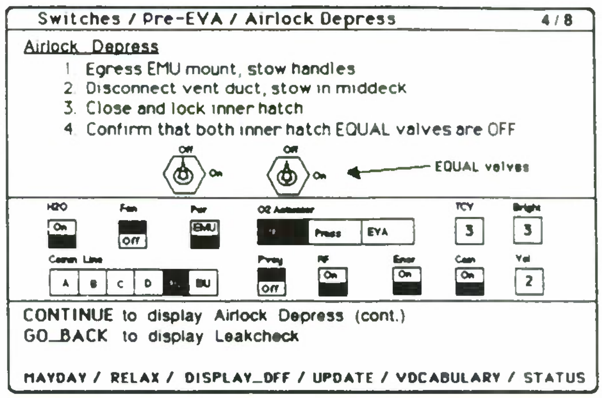


Figure 6. Airlock Depress Checklist (page 1)

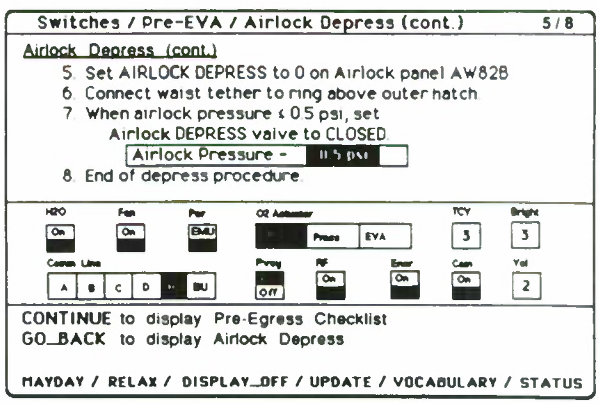


Figure 7. Airlock Depress Checklist (page 2)

(5) Once the simulated depressurization to vacuum is accomplished, the crewmember will "CONTINUE" to the Pre-Egress Checklist (see Figure 8). (Note that the crewmember may switch several controls with a single voice command.) Once he has completed this checklist, he will "CONTINUE" to the Egress Checklist (Figure 9). The Pre-Egress and Egress procedures lead the crewmember safely out of the airlock to perform the actual EVA. After the final page of this checklist has been read, an "EXIT" command will display the Normal EMU Status page for use during the EVA.

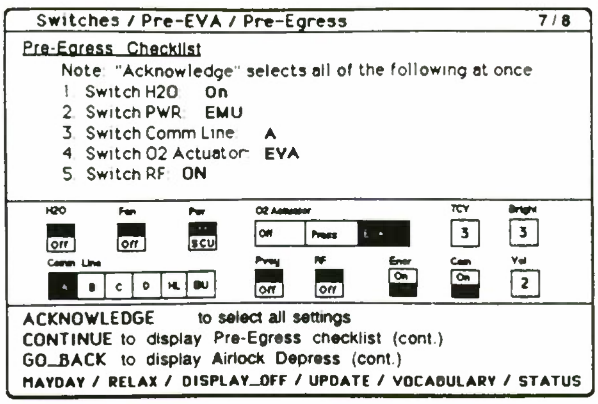


Figure 8. Pre-Egress Checklist

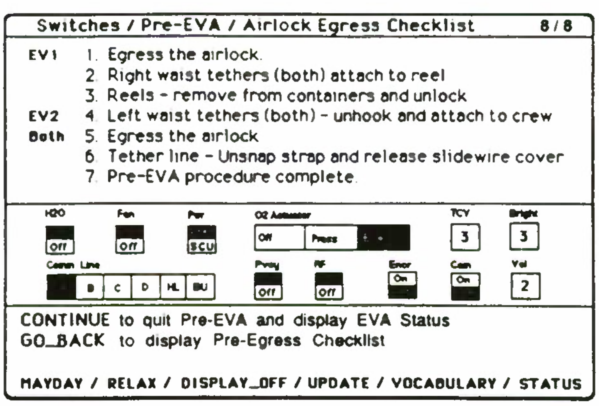


Figure 9. Egress Checklist

It should be noted that during the pre-EVA simulation (as will be in the flight system) other functions which include display on/off and brightness controls, for example, will be available to the user.

EVA

The actual EVA, by definition, begins when the crewmember leaves the airlock. This begins the operational portion of the astronaut's venture. Nearly all of the HMD's functions are available to the EVA crewmember at this time. During this period, the EVA astronaut may view EMU status information, access general EVA operational information, control EMU switches by voice, monitor transmissions from the astronaut's personal TV camera, and obtain data directly from the Space Station's main computers via radio link. The simulated functions are described below along with a description of other available functions.

EMU Status - The EMU Status file is the base state of the simulation program. Each page is split into two fields (excluding the header) as shown in Figure 10. The top field displays specific EMU parameters and consumable values in two formats: on a quick-read bar graph format and actual (or calculated) values. This method allows the astronaut to quickly and accurately assess the performance of the EMU. The lower field lists valid available functions.

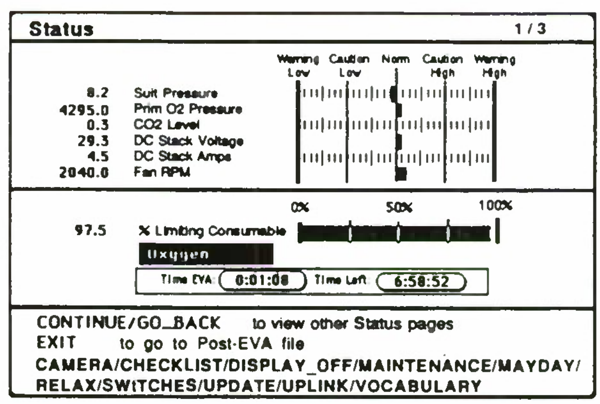


Figure 10. Normal EVA Status

EVA Operational Information - This subroutine allows the astronaut to select and read various data files from a main menu by voice commands. By saying "MAINTENANCE," the user obtains the menu as shown in Figure 11. This file and the "CHECKLIST" file

provide information about maintenance procedures, equipment operation, and contingency or troubleshooting flow charts, among other information subfiles in text and graphical formats. Once accessed, the user may page back and forth through the file. As an example, if "THREE" were selected, the crewmember would retrieve the "PAM Large Sunshield" file, which details a procedure for installing "sunglasses" on the Space Shuttle's windows. Page one of the "PAM Large Sunshield" file lists the tools the EVA crewmember must collect to perform the job (see Figure 12). From here, the astronaut may "CONTINUE" to obtain the next page of information. An "EXIT" within a file would bring the user back to the main menu. Another "EXIT" would bring the user to the base state, the first Status page.

EMU Switch Manipulation - (This procedure is identical to that described in the pre-EVA section.)

EMU Camera Monitoring - Shuttle EVA astronauts are not capable of monitoring TV cameras. They depend on ground or Shuttle communication for feedback on pointing the camera. On the other hand, with the Station EMU, an astronaut can say "CAMERA" and receive live video from his TV camera directly on his HMD. Currently the simulation program only displays still graphics to the user. An enhancement to the simulation program will soon be added to allow the crewmember to move and position a camera by voice.

Accessing Information From The Station - From the Station's central computer, the EVA astronaut may receive unlimited amounts of information via a radio frequency link to the crewmember. This information, in the form of files, will be accessed by the EVA crewmember much like the operational information described above. These files may have been produced prior to launch, prior to the EVA by that astronaut, or by someone on the ground to be transmitted, real-time, to an EVA astronaut. A procedure to simulate this function is currently being developed.

```
Maintenance / Main Menu                                  1 / 1

   1  Trash Disposal                                 (ONE)
   2  Standard Sunshield Manual Opening              (TWO)
   3  PAM Large Sunshield                            (THREE)
   4  MMU Donning Procedure                          (FOUR)

ONE-FOUR  to select file
EXIT      to return to status

MAYDAY / RELAX / CLEAR DISPLAY
```

Figure 11. EMU Maintenance File

```
Maintenance / PAM Large Sunshield / Tools                1 / 2
   Tool box
      1. Probe & hammer
      2. Trash container
      3. 3/8" drive ratchet
      4. Adj wrist tether
      5. Cable cutters

   Airlock
      1. IFM tool (7/16")
      2. PAM tool (Spreader)
      3. IFM driver handle with 3/8" socket

CONTINUE  to display next page
EXIT      to return to main menu

MAYDAY / RELAX / CLEAR DISPLAY
```

Figure 12. PAM Large Sunfield File

Other Functions - Other functions are also available to the crewmember while receipt of alarms with corrective action, notifying Station astronauts of anomalous conditions (which includes an option for sending them EMU transmitted data), a complete recognizer vocabulary menu, and updating voice templates by voice.

Post-EVA - To complete the EVA, the astronaut enters the airlock and connects his umbilical line to the EMU. Post-EVA instructions are provided to the crewmember via the HMD and are basically a reverse sequence of the pre-EVA set of events, concluding with a power-down and doff of the EMU.

CONCLUSION

Both voice recognition and helmet-mounted display technologies can improve the productivity of workers in space by potentially reducing the time, risk, and cost involved in performing manned extravehicular activity. NASA has recognized this potential and is currently developing a voice-controlled information system for Space Station EVA. Two bench-model helmet-mounted displays and an EVA simulation program have been developed to demonstrate the functionality and practicality of the system.

USE OF INFRARED TELEMETRY AS PART OF A NONINTRUSIVE INFLIGHT DATA COLLECTION SYSTEM TO COLLECT HUMAN FACTORS DATA

Angelo Micocci
Lockheed Engineering and Sciences Company

Research directed by Francis Mount, Flight Crew Support Division, NASA JSC.

INTRODUCTION

The objective of this paper is to present a methodology and rationale for development of a Nonintrusive Inflight Data Collection System (NIDCS) to collect Human Factors (HF) data during a space mission. These data will enable the research team to identify and resolve issues.

This paper will present the background and history of the NIDCS, the methodology and techniques employed versus those in current use on Earth, initial results of the effort, including a brief description of the equipment, and, finally, a discussion of the scientific importance and possible future applications of this system elsewhere.

The schema for the NIDCS includes a collection of three types of data, behavioral, physiological, and biomechanical. These will be collected using videotape of crewmembers' activities, bioelectric signal measurement, and measurement of kinematics and kinetics, respectively. This paper will focus on the second type of data, physiological activity as determined by changes in bioelectric potentials as crewmembers perform daily assignments.

BACKGROUND AND HISTORY

Before proceeding, it is necessary to define nonintrusive data collection. The strictest definition of such a system is one in which the subject is unaware of the recording apparatus, i.e., has no knowledge of the fact that data are being collected, etc. Given training procedures, awareness of video cameras, audio report/debriefing taping on Earth, etc., the crew will be aware that data are being recorded. Therefore, the alternative solution to data collection is to employ a system that will not (or only minimally) disrupt routine, ongoing activities (Callaway, 1975). In any case, nonintrusive data collection should not be confused with "noninvasive" data collection, a method frequently used in collection of biomedical data. While nonintrusive methods are noninvasive, the reverse is not true. This distinction excludes collection of blood or other activities where catheterization or collection of other data which alter crew activities, etc., is necessary. We can now proceed to the history of the development of an NIDCS.

A review of the literature from other projects using nonintrusive data collection methods provides a basis for understanding the reasons for the application of such methods, how other methods were employed, the degree of their success, and what was learned from them. The literature search focused on situations likely to be encountered during space flight.

Literature was reviewed from the SEALAB (Radloff, 1966), TEKTITE (Nowliss, 1972), and SKYLAB (LaFevers, NASA JSC) missions. The schedule for development of an NIDCS precluded review of data from submarines, due in part to time limitations and because most submarine data have been collected and evaluated for other purposes (contamination monitoring/control, etc.) and are not structured for application to HF issues. Because data collected on the referenced missions focused on behavioral rather than physiological data, the detailed findings on these projects will not be discussed here. Researchers found that the most expeditious way to collect (behavioral) data was through simple, passive observation of video and through the use of questionnaires and post-mission debriefings. It is reasonable to expect

that passive recording of data to yield results can be applied to physiological and biomechanical data as well.

METHODOLOGY AND TECHNIQUES

As mentioned in the introduction, there are three types of data to be collected using the NIDCS during space missions.

First, there are the behavioral data wherein the primary method used for collection is videotapes. Except for those instances where arrangements were prescheduled, researchers were able to see only portions of activities which might be of interest to their disciplines. Camera angles were based on factors other than those which the researcher might select, duration of taping was not optimum, detailed verbal explanations were not forthcoming, etc. The best to be hoped for was that the crew member(s) could provide sufficient details during debriefings, either in flight or postflight. While satisfactory, it is not optimum for determination and resolution of issues.

One of the objectives of this project is to provide input and direction for the program during the early planning stages to ensure that Human Factors-oriented tasks are part of the mission; to attend training and planning sessions involving crewmembers to ensure that training includes performance of these tasks when required; and to specify video camera locations, angles, etc., to ensure that adequate coverage exists for evaluation. Videotapes will be supplemented by the information obtained during debriefings, through the use of questionnaires, etc.

Secondly, there are the biomechanical data which will provide information concerning kinetics and kinematics during performance of space-oriented tasks. Current planning is that these data will be collected in conjunction with another experiment dealing with human force capacity in space. It is anticipated that some of these data will be collected using video cameras, while most will be obtained with the use of sensors and instruments such as transducers, force plates, accelerometers, strain gauges, etc.

Finally, there are the physiological data. Much of modern research has focused on clinical applications, diagnostic purposes to develop prosthetic devices, etc. Various studies have been performed to measure muscle fatigue, per se. Only recently (in the past 25-30 years) has research been done toward interpretation of these data as indicators of human performance (sports medicine, etc.) and/or operator alertness.

Some work has been done in manufacturing facilities, some in academic laboratories (Evoked Potentials Response-ERP) and some by HF engineers involved with the Space Program. These studies have been conducted to measure operator stress and alertness during the performance of assigned tasks. With regard to space, studies have focused on measuring electromyograms (EMGs) collected from subjects operating under both shirtsleeve and pressure suited conditions while performing simulated space vehicle control tasks. Some studies have been performed using electrocardiograms (ECGs) or electro-oculograms (EOGs) as measures of stress and performance, etc.

This research has been performed in settings where the subject is connected to the recording device(s) through a system of cables or "hard wires." In a simulated space vehicle control environment, the subject might exert forces against some controller device, track a target, etc. This approach is suitable and works well for Earthbound experiments; however, it presents problems during a space mission. On Earth, gravity acts on the interconnecting cables and keeps them in a cohesive bundle. This minimizes interference with task performance. In most Earth testing situations, cables dangling from the subject present no problems. Such is not the case during space flight. Feedback from videotapes, postflight debriefings, and other sources indicate that the interconnecting cables tend to float and interfere with task performance. This is true of audio system cables as well as cables interconnecting experiments, crewmembers,

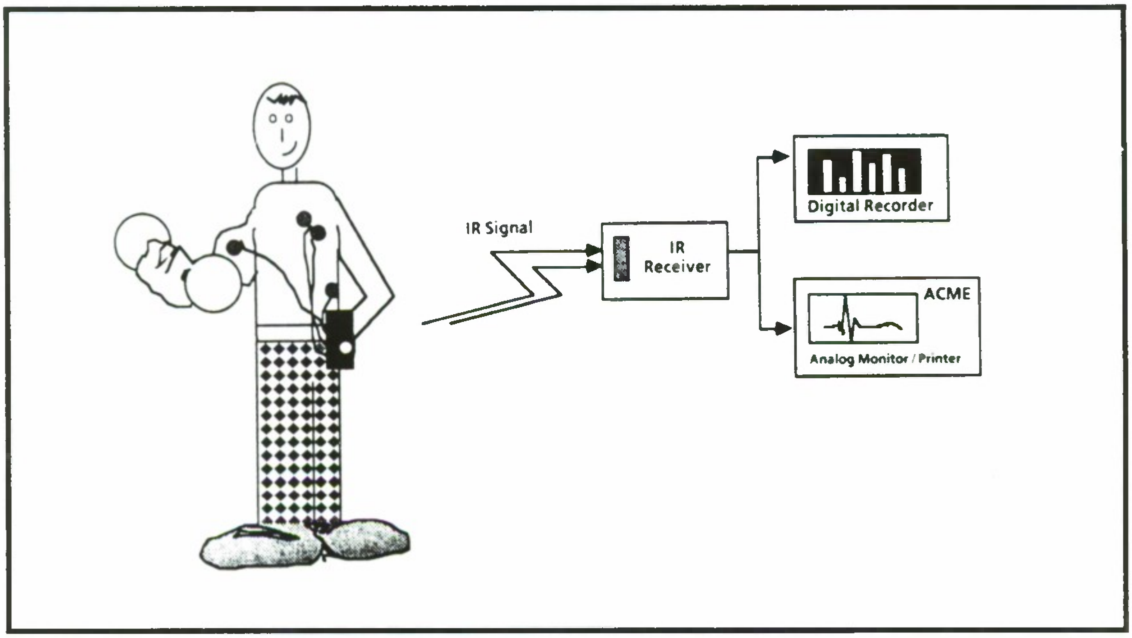


Figure 1. Paradigm of Telemetry System

etc. Much time is spent sweeping these cables out of the way. Shortening of cables provides an unsatisfactory solution, since this restricts crewmembers' mobility and range. These types of problems, concomitant with the HF experimenter's desire to collect data under normal conditions, make the use of hardwiring undesirable for data collection during space missions. With these considerations in mind, the NIDCS was devised.

EQUIPMENT

Meetings were held to discuss the viability and feasibility for developing a "noninterference" (telemetry) system for transmission of physiological signals. Infrared was selected as the signal carrier because the RF channels are assigned for flight data or other critical inflight parameters. There are few channels available for other uses. Also, IR tends to reflect off certain surfaces and increases the probability of detection.

A system was developed by modifying existing equipment. This modified system contained both a transmitter and a receiver unit with an "end-to-end" response of 6 kHz. It was felt that this bandwidth would allow the systems to handle 3 channels of physiological data. This portion of the NIDCS is referred to as the IPDL (Infrared Physiological Data Link).

Further modifications were performed on the unit to increase wearing comfort, to adapt the unit to physiological data handling, etc. A "breadboard" model was developed and tested. Testing was accomplished through the use of commercially available electrodes attached to a subject and to specifically developed signal conditioners through a (body-worn) cable which then attached to the IPDL transmitter (Figure 1). The subject then walked, flexed muscles, and simulated task performance. Data were recorded on a 4-channel strip chart recorder where one channel was used as the event marker. None of the data collected during testing were analyzed since they had not been collected using strict scientific standards or practices. They were used only to indicate the presence of bioelectric potentials. System reliability was determined through recording signals on the strip chart, both through the IPDL system and hardwiring the subject to the recorder. Wave forms were then compared for

distortion or any other changes which might have occurred and could be attributed to the IPDL.

The system operated satisfactorily except for interference (crosstalk) between the center channel and both end channels. As a result, plans were revised from development of a channel unit to a 2-channel system. Using the same approach as described above, a 2-channel prototype unit was developed, tested, and found to be satisfactory.

Experiments were designed, and a pilot study was run. Data were collected through the use of an instrument-quality tape recorder. Signals were then played into a personal computer for data storage and waveform and statistical analyses. The results of that study are being presented separately (by F. E. Mount, NASA JSC, under whose auspices this project was funded and directed) and will not be reported herein.

FUTURE APPLICATIONS

The importance of this effort is that a means has been developed which allows data collection while reducing, if not eliminating, certain types of experimental bias, such as the guinea pig effect, role selection, response sets, and measurement as an independent variable. Further, experimental error created by the investigator (subject induced bias due to sex, race, etc.) can be eliminated, data collection can occur under everyday settings rather than a laboratory (which may affect some data), errors due to changes in procedures or instructions are eliminated, etc. Finally, data can be collected that might not be collected in any other way.

Future applications for this approach are seen in the fields of medicine where restriction of a patient's movement is undesirable and in other similar situations. This technique could be utilized to make patients with certain disorders more independent by allowing them to be away from a laboratory or medical environment while still allowing for monitoring of homeostatic functions on a continuous basis.

Other research areas which need to be addressed are development of better, more natural and more comfortable electrodes, digitizing signals at the transmitting unit to compress data for storage, etc. This is only a short, and by no means exhaustive, list of areas for future research.

REFERENCES

Due to space limitations, the following is an abbreviated list of references.

1. Basmajian, J. V. *Muscles Alive* (Williams and Wilkins Co., 1974).

2. Callaway, E. *Brain Electrical Potentials and Individual Psychological Differences*. (Grune and Stratton, 1975).

3. Gale, A. *Some EEG Correlates of Sustained Attention*. Vigilance (Plenum Press, 1977).

4. Gould, J. D. Eye Movements During Visual Search. IBM Research Report RC 2680 (Yorktown Heights, NY, 1969).

5. Kubis, J. F., et al. Time and Work on Skylab Missions 2, 3 and 4. Biomedical Results From Skylab. NASA SP-377 (1977).

6. LaFevers, E. A., and Hursta, W. Final Report: Lower Spectral Density Analysis of the SkyLab 3 Data. NASA Report (JSC).

7. LeVeau, B. *Biomechanics of Human Motion* (W. B. Saunders, 1977).

8. McKay, R. S. *Biomedical Telemetry* (John Wiley, 1965).

9. NASA Workshop on Advances in NASA-Relevant, Minimally Invasive Instrumentation. (Pacific Grove, CA, April 1984).

10. Nowliss, D. P., Wortz, E.C., and Waters, H. TEKTITE II, Habitability Research Program. Airsearch Mfg. Co. Report. 7-6192-1 (January 1972).

11. Radloff, R., and Helmrich, R. *Groups Under Stress: Psychological Research in SEALAB II* (Appleton-Century-Crofts, 1968).

12. Weaver, C. S. Wearable Blood Pressure and ECG Recording System. Stanford Research Institute Interim Report (February 1976).

PERFORMING SPEECH RECOGNITION RESEARCH WITH HYPERCARD

Chip Shepherd
Lockheed Engineering and Sciences Company

Research directed by Barbara Woolford, Flight Crew Support Division, NASA JSC.

PURPOSE

The purpose of this paper is to describe a HyperCard-based system for performing speech recognition research and to instruct Human Factors professionals on how to use the system to obtain detailed data about the user interface of a prototype speech recognition application.

BACKGROUND

The development of the first Macintosh-based speech recognizer (Voice Navigator by Articulate Systems, Inc.) has enabled engineers at the NASA Johnson Space Center (JSC) to develop rapid-prototype speech recognition interfaces for space applications with HyperCard, an information management software package. A layout of the required hardware is presented in Figure 1.

Just like most speech recognizers, the Voice Navigator will allow a user to define a unique vocabulary, assign the computer action(s) to be associated with each vocabulary word, and record a personalized voice pattern for each word. The Navigator goes a step farther than most recognizers, however, because it allows access to the various recognition parameters generated within the machine while it is in an operating mode.

To obtain data on speech recognition parameters while the unit is being operated, engineers have taken advantage of the fact that HyperCard's command language may be expanded through the implementation of executable commands (XCMDs), which are user-created subroutines written in assembly or C languages.

In JSC's speech recognition research, three types of recognition data were required each time a command was spoken: the recognized word, the loudness of the speech, and the "confidence score," which is an internally calculated measure of how closely the spoken word compares to a pre-recorded sample; the score is expressed as a number from 1 to 100, with 100 being a perfect match. Using specifications from JSC contractor engineers, special HyperCard XCMDs were developed by Articulate Systems to obtain speech recognition data and assist in JSC's research applications. By integrating the XCMDs into experiments, project engineers have learned how to use the XCMDs to perform quantitative speech recognition research.

XCMD DESCRIPTIONS

Articulate Systems included twelve XCMDs in a HyperCard stack called VoiceTalk, created to assist with the completion of the speech recognition research. Seven of the twelve XCMDs were designed to control necessary file management tasks. The other five have been applied in conducting research and are described below.

"Vocabulary" (name of vocabulary file) - Writes the active vocabulary list into the local HyperCard variable 'it.' The subject and/or researcher may be informed at all times which vocabulary set the recognizer is listening for.

"Collect" (vocabulary word) - Collects new voice samples of the specified vocabulary word. Thus, the researcher may record new voice samples just before, or even during, an experiment session.

"Macro" (vocabulary word) - This feature returns the command, or macro, associated with the specified vocabulary word. The macro is

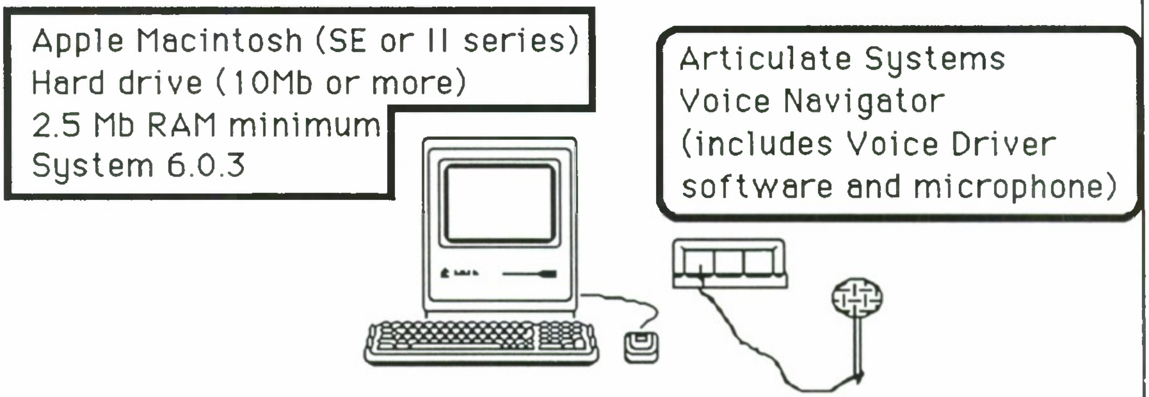


Figure 1. Recommended minimum hardware for HyperCard-based speech recognition research system.

usually a string of text, but can also be a mouse click or a series of HyperTalk commands, complete with returns, tabs, and linefeeds.

"Listen" - Prepares the recognizer to accept speech input.

"Recognize" - When an utterance is detected, this command returns the name of the recognized word. In addition, several associated speech recognition parameters are stored in reserved variables. The most valuable of these associated parameters is the confidence score, which is stored in the reserved variable "Confidence." The loudness of the speech is reported in decibels in the reserved variable "Amplitude."

GETTING THE SYSTEM TO WORK FOR YOU

The present Voice Navigator, which is commercially available, does not include the HyperCard XCMDs as part of its standard package. This is because the XCMDs were created while the Voice Navigator was a prototype unit. However, the XCMDs may be easily obtained from Articulate Systems and will operate with available hardware after a few minor software adjustments are made by the user.

To utilize the XCMDs in HyperCard applications, the researcher should follow the procedures outlined below. These procedures will work with Voice Navigator System 1.0.1:

(1) Obtain a Voice Navigator from Articulate Systems, specifically requesting VoiceTalk.

(2) After you have connected the Voice Navigator and installed its software system files into the Macintosh as described in the Voice Navigator User's Manual, use the "Duplicate" command to make copies of the "Voice Control" and "Voice Driver" files that reside in the System folder. You should now have two new files named "Copy of Voice Control" and "Copy of Voice Driver."

(3) Change the name of the file "Copy of Voice Control" to "Dragon" and "Copy of Voice Driver" to "Dragon.LOD." You must do this to get the XCMDs to operate because, as stated before, the XCMDs were created when the Voice Navigator was in the prototype stage and, at that stage, the driver files were named "Dragon" and "Dragon.LOD." The Macintosh must be rebooted before the new files can be accessed, so you may as well do so before continuing to the next step.

(4) Install the HyperCard XCMDs onto whichever HyperCard stack(s) you will be using for speech recognition applications. You

may accomplish this with either the ResEdit program or the "install" feature provided with the VoiceTalk stack.

(5) Now that the XCMDs are installed, you may include them into any HyperTalk script on the stack. Examples of how to integrate the commands into HyperTalk programs are included with the VoiceTalk package.

Articulate Systems reports that improved versions of the HyperCard stacks will be available in the near future, but the procedures just described should enable presently available components to function.

DISCUSSION

Using the XCMDs created by Articulate Systems, engineers at JSC have been able to extract important information about the speech recognition application unobtrusively in real time as the user operates the application. Information about the recognized word, its confidence score, the loudness of the speech, and the elapsed time may be recorded in an invisible background data field that is stored and analyzed after the user has completed the session. With this system, the collection of the speech parameter data has not affected the user interface and has not added noticeable time delays to the execution of the interface programming.

EXAMPLE APPLICATION

The described system has been used extensively at JSC in examining speech as a means of controlling a computer display in the extravehicular environment. Future space suits may be equipped with a Helmet-Mounted Display (HMD), capable of providing substantial amounts of text, graphics, and video information. Controlled by spoken commands, this system would provide hands-free access to information for the suited astronaut.

A test was conducted with a prototype HMD system at NASA-JSC by having subjects use the HMD-based information system to construct an electronic circuit (Shepherd, 1989). The speech recognizer was active at all times and four types of data were collected each time a word was recognized: the word, the confidence score, the amplitude, and the elapsed time for the session (see Figure 2). A segment of the HyperTalk script used to perform these functions is included at the end of this paper (see EXAMPLE STACK SCRIPT).

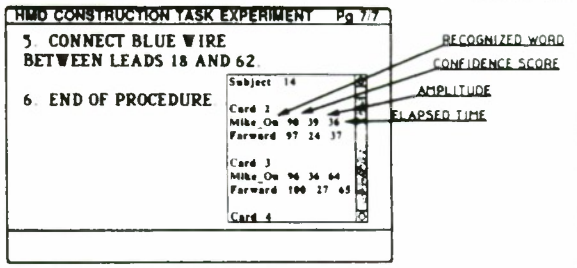


Figure 3. HyperCard screen from HMD construction task experiment with data screen shown.

A list of the recognized words was correlated with the elapsed times and compared to the videotape of the experiment, enabling researchers to easily classify each entry as a correct recognition or an error. Upon further analysis of the errors, it was found that the most common type of error occurred when the recognizer had inexplicably registered a command from the subject's conversational speech.

The confidence scores of the errors were analyzed and compared to those of the correct recognitions. The patterns were very different. While almost all of the correct recognitions had scores above 70, the errors were scattered throughout the range of scores, indicating that the errors had occurred randomly.

The amplitudes of the errors were analyzed and compared to those of the correct recognitions to see if the errors were said more loudly or more quietly than commands. No significant differences were found, which indicated that the errors were said just as loudly as the commands.

It should be noted that the XCMDs have not only been useful in collecting the data but are also helping to improve the interface itself. Because of the difference in confidence score patterns between correct recognitions and errors, system performance has been increased by setting a confidence score threshold in the system (i.e., each time a word is recognized, the corresponding confidence score, which is generated by the XCMD, is checked and, if it is not 70 or above, the command is ignored). With this threshold in place, many errors are screened out while almost no correct recognitions are affected. A similar threshold for the amplitude could have been put into place if a difference in the patterns had been found.

The described system has been beneficial in studying speech control for space applications, but it can also be employed in evaluating prototype interfaces in any of the leading speech recognition fields, including medicine, defense, products for the handicapped, and consumer systems.

EXAMPLE STACK SCRIPT

Explanation: The first script basically instructs the recognizer to keep listening until it hears a vocabulary word. Once a vocabulary word registers, a second script would be triggered. The script for "Mike_On" is provided as a sample. This script "activated" the microphone to accept a page forward/backward command. However, if the confidence score was not 30 or more, the script instructs the recognizer to ignore the command it just heard and to start listening for another.

```
On OpenCard
  global RecognizedWord, StartTime
  put the seconds into StartTime
  put empty into RecognizedWord
  listen
  repeat until the mouseclick
    put recognize() into RecognizedWord
    if RecognizedWord is not empty then
      get macro(RecognizedWord)
      do it
    end if
  end repeat
end OpenCard
```

```
on Mike_On
  global RecognizedWord, StartTime,
Confidence
  global Amplitude
  if Confidence > 30 then
    put RecognizedWord &&
Confidence&&Amplitude¬
    && ((the seconds) - StartTime) & return¬
    after bkgnd field "Recording"
  else
    put 3 into dsd_state :(i.e., ignore the
command)
  end if
end Mike_On
```


ACKNOWLEDGEMENTS

The author would like to thank Barbara Woolford and José Marmolejo of NASA-JSC and Tim Morgan of Articulate Systems for their support in this research.

References

1. Articulate Systems, "Voice Navigator User's Manual," 1989.

2. Cohen, A. D. "Rapid Prototype Development with HyperCard," Workshop at the 33rd Annual Meeting of the Human Factors Society, October 1989.

3. Shepherd, C. K., Jr. "Human Factors Investigation of a Speech-Controlled Bench-Model Extra-vehicular Mobility Unit Information System in a Construction Task," NASA document no. JSC-23632, Lockheed Engineering and Sciences Company document no. LESC-26799, May 1989.

4. Shepherd, C. K., Jr. "The Helmet-Mounted Display as a Tool to Increase Productivity During Space Station Extravehicular Activity," Proceedings of the 32nd Annual Meeting of the Human Factors Society, October 1988.

5. Weimer, J. "HyperCard for Human Factors Applications," Workshop at the 32nd Annual Meeting of the Human Factors Society, October 1988.

ILLUMINATION REQUIREMENTS FOR OPERATING A SPACE REMOTE MANIPULATOR

George O. Chandlee and Randy L. Smith
Lockheed Engineering and Sciences Company

Charles D. Wheelwright
NASA Lyndon B. Johnson Space Center

INTRODUCTION

Critical issues and requirements involved in illuminating remote manipulator operations in space help establish engineering designs for these manipulators. A remote manipulator is defined as any mechanical device that is controlled indirectly or from a distance by a human operator for the purpose of performing potentially dangerous or hazardous tasks to increase safety, reliability, and efficiency. Future space flights will rely on remote manipulators for a variety of tasks including satellite repair and servicing, structural assembly, data collection and analysis, and performance of contingency tasks. Carefully designed illumination of these manipulators will assure that these tasks will be completed efficiently and successfully.

Studies concerning the influence of illumination on operation of a remote manipulator are few. Available results show that illumination can influence how successfully a human operates a remote manipulator. Previous studies have indicated that illumination should be in the range of 400-600 footcandles, the currently recommended range for fine to medium assembly work tasks [1, 2, 3, and 4]. However, on-orbit operations have demonstrated that effective operation can be performed under illumination levels between 100-500 footcandles provided that specularity and glare are minimized. Increasingly complex and finer work tasks would necessarily require an increase in illumination [1, 2, and 3]. Brightness should not exceed 300-450 footlamberts (roughly equivalent to a zenith sky or slightly brighter [3]) so as to eliminate glare and possible blinding of the operator. The reflectance percentage of the manipulator and target should be about 75%, if possible [5]. The beneficial aspect of high specularity is that detection distances may be as great as 5 miles during rendezvous operations [5]. Reflectance is an especially critical illumination-related parameter because a target object and a manipulator should be visible to a human operator from distances of at least 30-40 meters.

Remote manipulators and their task and target objects are operated optimally when direct and indirect glare is eliminated and lighting tends to be diffuse. Several ways exist to reduce glare: position light sources outside the operator's line of sight, use low-intensity light sources, increase luminance of the area around glare sources (direct glare), position light sources so a minimum amount of light is directed toward the eyes to prevent frontal and side blinding, and use luminaires with diffusing or polarizing lenses. Reflective surfaces should diffuse the light (flat paint, nongloss paper, and textured finishes). Illuminance levels should be kept as low as possible and use indirect lighting [2, 4]. A neutral density filter with a transmission of 20 percent can help reduce general reflected glare, except for the very harsh specular type, to an acceptable level [6].

Design of lamp lenses is an important factor to be considered in establishing lighting requirements. Lenses distort light distribution and can otherwise alter viewing conditions and may contribute to poor operator performance. Previous studies indicate that a planar-planar lens is probably the best because proper illumination levels will be maintained and contrast ratios are adequate [7]. Planar-planar

lenses are currently used in Space Shuttle cargo bay lights.

MAJOR ISSUES

Literature surveys, evaluation of published experimental results, and consideration of requirements in Earth-based operations suggest that seven major issues may be identified as major contributors to establishing illumination requirements for a remote manipulator in space. These seven issues are listed and discussed below.

(1) *Sun angles and their influence on reflectance/viewing characteristics of remote manipulators.* Wheelwright [5, 6] has shown experimentally, with the use of scale models, the effects that sun angles have on the operation of the remote manipulator system on the Space Shuttle. Effects of sun angles depend largely on the viewing configuration. Remote manipulator teleoperations should avoid direct sun viewing to prevent blinding of the operator. Specular glare, veiling lumens, and extremely bright areas occur at various sun angles. Although objectionable, most on-orbit operations can still be completed. Reorientation of the viewing angle by no more than 5 degrees will, however, produce more favorable lighting conditions [6].

(2) *Reflectance properties of the manipulators under solar illumination and artificial lighting.* Comparison of the reflectance properties of manipulators under solar and artificial lighting is important to establish when and how each lighting regime can be used to advantage for optimum performance, to determine proper artificial light sources, and to determine what special filters might be necessary to enhance recognition and minimize specularity. Because solar light is collimated, special considerations may involve reflectance characteristics so as to minimize deleterious edge effects encountered in on-orbit operations.

(3) *Recognition and reflectance properties of task structures and target objects.* Recognition of task structures and target objects is important in order to optimize operations and to reduce hazards associated with incorrect identification. Target-background contrast ratios of at least 0.6 should be used to attain optimum size discrimination in two-target tasks. Size discrimination performance depends on target-background contrast. With contrast ratios of at least 0.6, the linear dimension size discrimination is on the order of 0.10. A reduced contrast ratio of 0.125, however, raises the threshold value to 0.30. Brightness discrimination between two targets is enhanced for contrast values of 0.25 or greater [8]. Some tasks may involve recognition of alpha-numeric characters. Character density, character contrast, viewing distance, and monitor size are some of the variables that can affect correct identification of alphanumeric characters [8].

(4) *Search and rendezvous requirements using running lights of a free-flying remote manipulator.* Establishing search and rendezvous requirements is critical because docking with a remote, free-flying manipulator will be an essential activity in on-orbit operations. Remote manipulator docking and operation require that the operator be able to acquire depth and range information from the visual system. Range estimation depends on target size, brightness, and contrast [8] Configuration of running lights is important because recognition of form and orientation of axes will be critical for successful rendezvous and docking operations.

(5) *Tracking and recognition of remote manipulators by direct vision and monitor viewing.* The issue of direct vision and TV monitor viewing remains unresolved. Evidently, each viewing method has advantages under certain conditions. Viewing is the primary form of feedback to the operator regarding manipulator position, orientation, and rate of movement. The illumination system and viewing system are interdependent and together result in operator perception of manipulator motion.

(6) *The influence of light intensity, position, and type on operator performance.* A few studies have measured how operator performance is influenced by light intensity, position, and type. Operator physical and

mental workload may be dramatically affected by these parameters. Onboard lighting is effective for close-in illumination of shapes and spaces hidden in deep shadow. A variety of lights differing in illumination output, power consumption, spectral specularity, beam width, and efficiency will probably be necessary for on-orbit operations. Results suggest that performance is best maximized by tailoring light intensity, position, and type to the specific task [9].

(7) *The effects of shadow patterns on operator interpretation and performance.* The interpretation of shadow patterns could have a significant effect on operator performance but just exactly what these effects are remains to be determined. Local task-specific lighting may be necessary to overcome some of the problems associated with shadow patterns. Total elimination of shadows appears impossible, however, and more research is needed to determine cognitive processes involved in shadow interpretation.

All seven issues must be resolved in the context of realistically achievable physical conditions in space. Perhaps two of the most limiting conditions will be the power availability (the power requirements for the use of remote manipulators) and thermal conditions. Power is a necessary, but limited commodity in a space environment and will be a restrictive factor in remote manipulator operations. Illumination designs and hardware must account for potential problems in heat dissipation.

DISCUSSION

Identification, characterization, and analysis of each of the seven major issues will contribute to engineering design plans for a successful human operator-remote manipulator interface. An initial approach to determine the relative importance of each issue with respect to remote manipulator operation is to establish some critical visual activities and how they are related to each of the seven issues. A suggested relationship among the seven issues discussed in this paper and six critical visual activities as defined by Huggins, et al. [10] is displayed in Table 1. Critical visual activities were defined and studied by Huggins, et al. [10] in their evaluation of human visual performance for teleoperator tasks.

In this study, we define acuity as keeness of perception, discrimination as the ability to distinguish among objects, and recognition as the ability to identify an object. Using these definitions, we suggest which of the six critical visual activities are most likely to be affected by each of the seven issues. In no way is this intended to mean that some activities are not affected by all the issues; indeed, each issue will have some, albeit small in some instances, influence on each activity. Instead, we infer some critical visual activities to be more vulnerable to anticipated specifications defined by the issues than others. Expectations will most probably change as new data become available, so the suggested relationships given in Table 1 should not be considered as fixed. Rather, the given relationships are intended only as a guideline to help plan, evaluate, and interpret future studies and, possibly, designs.

Once each illumination parameter has been specified, continuing human factors engineering studies should evaluate the various kinds of mental models used by an operator. Mental models can be used to account for human interactions with a remote manipulator by helping to define the cognitive processes involved in human-remote manipulator interactions [11]. Definition of mental models used by an operator of a remote manipulator could help establish lighting arrangement and intensity, brightness and illumination requirements, and mental processes associated with shadow interpretation.

The major issues identified in this study may also be helpful in defining illumination requirements in other applications using indirect human operation and in creating optimum engineering designs for remote manipulators used in undersea tasks, assembly-line work, and in the nuclear industry.

TABLE 1

Summary of suggested relationships among the illumination issues discussed in this paper and critical visual activities evaluated by Huggins, et al. [10]. (See text for definition of terms and issues.) Each X shows what critical activities are most likely to be affected by the issues discussed in the text.

ILLUMINATION ISSUE	CRITICAL VISUAL ACTIVITY: Acuity Size	Size Estimation	Form Discrimination	Brightness Discrimination	Pattern Recognition	Depth Distance
Reflectance of remote manipulator	X		X	X		X
Solar vs. artificial lighting			X	X	X	
Reflectance of target	X	X		X		X
Effect of running lights			X			X
Direct vision vs. TV	X				X	X
Lighting parameters	X	X	X			X
Effect of shadows	X	X	X		X	X

SUMMARY

(1) Preliminary guidelines for illumination requirements in remote manipulator tasks in space are suggested: illumination should be in the range of 400-600 footcandles (although under some circumstances a range of 100-500 footcandles could suffice), brightness should not exceed 300-450 lamberts, reflectance of target/task objects should be about 75% (we suggest a range of between 50-75%), and an optimum contrast ratio between target and background is at least 0.6%.

(2) Seven major issues related to illumination of a remote manipulator in space are discussed: the influence of sun angles on reflectance/viewing characteristics of remote manipulators, reflectance properties of the remote manipulator under solar and artificial lighting, task/target object recognition and reflectance properties, rendezvous requirements, tracking and recognition of remote manipulators by direct vision and TV monitors, lighting parameters (such as intensity, position, and type), and the effect of shadow patterns on operator interpretation and performance.

(3) Critical visual activities, such as acuity, discrimination processes, and recognition tasks in the optimum operation of a remote manipulator, are known to be influenced by the illumination environment. Future research intended to measure and interpret fully the influence of illumination should use scale-model and full mockup testing to evaluate human operator performance under various illumination regimes.

ACKNOWLEDGEMENTS

We thank Lockheed Engineering and Management Services Company and the Lyndon B.

Johnson Space Center, National Aeronautics and Space Administration, for supporting this research under contract NAS 9-17900 and for granting the opportunity to publish our results.

REFERENCES

1. W. H. Cushman and B. Crist. Illumination, in: G. Salvendy (Ed.), *Handbook of Human Factors*, John Wiley and Sons, New York, 1987, pp. 670-695.

2. R. D. Huchingson, *New Horizons for Human Factors in Design*. McGraw-Hill Book Company. New York, 1981.

3. P. R. Boyce. Human Factors in Lighting, Macmillan, New York, 1981. W. F. Grether and C. Baker, Visual presentation of information. in: H.P. Van Cott and R. G. Kinkade (Eds.), *Human Engineering Guide to Equipment Design*. U.S. Government Printing Office, Washington. 1972, pp. 41-121.

4. W. F. Grether and C. Baker, Visual presentation of information, in: H.P. Van Cott and R. G. Kinkade (Eds.), *Human Engineering Guide to Equipment Design*, U.S. Government Printing Office, Washington. 1972, pp. 41-121.

5. C. D. Wheelwright, Illumination evaluation using 1/10 scale model of payload bay and payloads. Phase II. Payload bav artificial lighting evaluation, Note 77 EW 4, 1977.

6. C. D. Wheelwright, Illumination evaluation using 1/10 scale model of payload bay and payloads. Phase I. Solar illumination evaluation. Johnson Space Center, Internal Note 77 EW 2, 1977.

7. N. Shields, Jr. and D.E. Henderson, Earth Orbital Teleoperator System Evaluation: 1979-1980 Test Report, Essex Corporation, Report Number H-81-01, 1981.

8. N. Shields. Jr., Francis Piccione, M. Kirkpatrick III, and T. B. Malone. Human Operator Performance of Remotely Controlled Tasks: A Summary of Teleoperator Research Conducted at NASA's George C. Marshall Space Flight Center between 1971 and 1981, Essex Corporation, Report Number H-82-01, 1982.

9. C. D. Wheelwright, Teleoperator retrieval system: Lighting evaluation, Johnson Space Center, Internal Note 79 EW, 1979.

10. C. T. Huggins, T.B. Malone, and N.L. Shields, Jr., Evaluation of human operator visual performance capability for teleoperator missions, in: E. Heer (Ed.) *Remotely Manned Systems: Exploration and Operation in Space*. Proc. of the First National Conference. California Institute of Technology, September 13-15, 1972, 1973, pp. 337-350.

11. R. L. Smith and D. J. Gillan, *Human-telerobot interactions: Information, control, and mental models*. Proc. of the Thirty-first Ann. Mtg. Human Factors Society, Volume 2, 1987, pp. 806-810.

PREVIOUS EXPERIENCE IN MANNED SPACE FLIGHT: A SURVEY OF HUMAN FACTORS LESSONS LEARNED

George O. Chandlee
Lockheed Engineering and Sciences Company

Barbara Woolford
NASA Lyndon B. Johnson Space Center

INTRODUCTION

Previous experience in manned space flight programs can be used to compile a data base of human factors lessons learned for the purpose of developing aids in the future design of inhabited spacecraft. The objectives of this study are to gather information available from relevant sources, to develop a taxonomy of human factors data, and to produce a data base that can be used in the future for those people involved in the design of manned spacecraft operations. A study is currently underway at the Johnson Space Center with the objective of compiling, classifying, and summarizing relevant human factors data bearing on the lessons learned from previous manned space flights. The research reported here defines sources of data, methods for collection, and proposes a classification for human factors data that may be a model for other human factors disciplines.

PERSPECTIVE

Three major manned space programs have been conducted since the mid-1960s: the Apollo, Skylab, and Space Shuttle programs. Each program has contributed significant new data to the field of human factors and to gaining a greater understanding of how humans operate, function, behave, and adapt to the environment encountered in space. Because of various circumstances, including time constraints, human factors data collected during the past two decades of manned space flight have been transferred in a way such that the data remain scattered in various locations and do not reside in a single central location that is accessible to interested individuals, including those that might be involved in future advanced spacecraft design. Difficulties encountered in the systematic collection of past data are compounded because new data and technology appear frequently and must also be stored for easy use by appropriate individuals.

METHODS

The technique for data collection involves identifying information sources including technical reports, films, or video tapes, minutes of meetings, and records of in flight and postflight debriefings. A taxonomy is imposed on these data and the taxonomy may be a model for other human factors research activities. Data are transferred to an appropriate data archival/retrieval system that serves as a resource from which individuals involved in spacecraft design can draw relevant human factors data. Data sources include documents, published and unpublished technical reports, individuals, transcripts of meetings, audio and visual recording media, and computer-stored information, and may be supplemented with information acquired in interviews or from questionnaires. Systematic collection of data from these identified sources involves establishing a way of coding, tabulating, and cataloging the data before incorporating it into a data management application for use. Since human factors data are so diverse, a scheme for classifying data helps impose a meaningful structure on the data and renders the data more easily incorporated into an appropriate data base.

Commercially available software packages have been evaluated as candidate applications for the development of the computer-based retrieval system. These packages generally fall into two categories: data base management systems and hypertext-based text/graphics

handling tools. Data base management systems correspond to the traditional linear and hierarchical method of storing data. This form of data archival and retrieval system does not take advantage of complex interconnected Links between textual/graphical data and, as a result, data browsing can be cumbersome and slow. Hypertext (and hypermedia) systems allow for complex organization of data (text and graphics) by allowing machine-supported references from one data unit to another by taking advantage of the ability of a computer to perform interactive branching and dynamic display (Conklin, 1987). In this fashion, hypertext systems allow a user to jump from one data unit to another through links. Data browsing becomes simple and efficient.

PROPOSED CLASSIFICATION SCHEME

The need for a classification system of human factors data has been recognized for years (Melton and Briggs, 1960), yet attempts to produce such a classification have been few. The main reason appears to be the belief that such a task would require enormous amounts of time and effort because of the quantity of literature and data available. The classification proposed here relates to human-machine interaction in the context of manned space flight but some aspects should be applicable to other endeavors.

The classification proposed here builds upon a previous one used to systematically categorize Skylab man-machine data. The groupings are operationally based. The following 19 categories are suggested to classify human factors lessons learned in previous manned spaceflight programs: Architecture, Communications, Crew Activities, Environment, EVA-suited activities, Food Management, Garments, Housekeeping, Locomotion, Logistics management (including failure management and the logistics and procedures involved in coping with system anomalies), Maintenance (scheduled and unscheduled), Manual dexterity, Mobility/restraint, Off-duty activity, Personal hygiene, Personal equipment, Physiological data, Tool inventory, and Waste management. Within each of these categories are other, less broad and more specific categories. The number of categories may appear large, yet previous studies have had to use many categories. Meister and Mills (1971), for example, created 18 categories in their attempt to determine requirements for and the elements of a human performance reliability data base. The number of categories must be large, given the large number of activities encompassed by human-machine interaction. Other workers (Chiles, 1967; Christensen and Mills, 1967) have indicated the difficulties involved in establishing a taxonomy of human factors.

DISCUSSION

Previous studies attempting to classify human factors data (Fleishman, Kinkade, and Chambers, 1968; Chambers, 1969; and Meister and Mills, 1971) have relied less on operational categories and more on behavioral and performance criteria. Meister and Mills (1971), for example, developed a taxonomy based on functional behavior and established categories such as auditory perception, tactile perception, and motor behaviors. This taxonomy reflected the goal of the study: to develop a data base of human behavior (behavioral data acquired during actual experimentation) for predicting human-machine performance. The taxonomy presented here reflects a different purpose: to develop a data base of human operator experience (operational experience data acquired as the result of various activities in space) for the purpose of providing a source of data to be used in the future design of manned-spacecraft operations.

Evaluations of presently available text/graphics software applications suggest that certain criteria must be considered when a data archival and retrieval system such as this one is developed. One fundamental criterion is the degree to which the system successfully retrieves relevant articles. The precision ratio (Lancaster, 1968) is one way of measuring this success. The ratio, developed in the context of information theory, is defined as R/L x 100 where R is the number of relevant documents retrieved in a search and L is the total number

of documents retrieved in a search. A retrieval application should have a high precision (at least 80% or above) to prove useful. To achieve such reliability, design of the application software and the data itself are critical. Flexibility, nearly unlimited growth potential, and the ability to effectively handle increasingly complex links that are established within the network are some attributes a viable application should possess.

Results presented here have significance in the establishment and design of a Space Station, lunar base, and Martian colony. The methods developed in the collection and systematic archival of human factors data during the course of this study may also bear directly on questions of ways in which to systematically compile and characterize human factors data in other areas of research including design of aircraft, automobiles, manned sea-faring vessels, and other similar activities. Research in human factors engineering has escalated in recent years and tremendous amounts of data are being generated (see, for example, Huchingson, 1981 or Woodson, 1981). Appropriate archival and retrieval systems will need to be developed to store this data.

Results of this work could have at least three possible applications: (1) other workers with large data sets might be able to use the collection methods developed in this study, (2) the taxonomy of human factors data developed will be applicable to other human factors research and might be used in instances where large volumes of unsystematically collected data exist, and (3) future research in the definition and design of human factors requirements may be able to benefit from the methods, taxonomy, and data organization techniques developed in this study.

ACKNOWLEDGEMENTS

We wish to thank the Lockheed Engineering and Sciences Company and the Lyndon B. Johnson Space Center, National Aeronautics and Space Administration, for supporting this research under contract NAS 9 - 17900 and for granting the opportunity to publish the results.

REFERENCES

1. Chambers, A. N. (1969). Development of a taxonomy of human performance: A heuristic model for the development of classification systems. Report 4A, American Institutes for Research. October, 1969. Silver Spring, MD.

2. Chiles, W. D. (1967). Methodology in the assessment of complex performance -- Discussion and conclusions. Human Factors, 9, 385-392.

3. Christensen, J. M., and Mills, R. G. (1967). What does the operator do in complex systems? Human Factors, 9, 329-340.

4. Conklin, J. (1987). A survey of Hypertext. MCC Technical Report. Number STP-356-86, Rev. 2.70.

5. Fleishman, E.,A., Kinkade, R.G., and Chambers, A.N. (1968). Development of a taxonomy of human performance: A review of the first year's progress. Technical Progress Report 1, American Institutes for Research, 59.

6. Huchingson, R. D. (1981). New horizons for human factors in design. New York: McGraw-Hill Book Company.

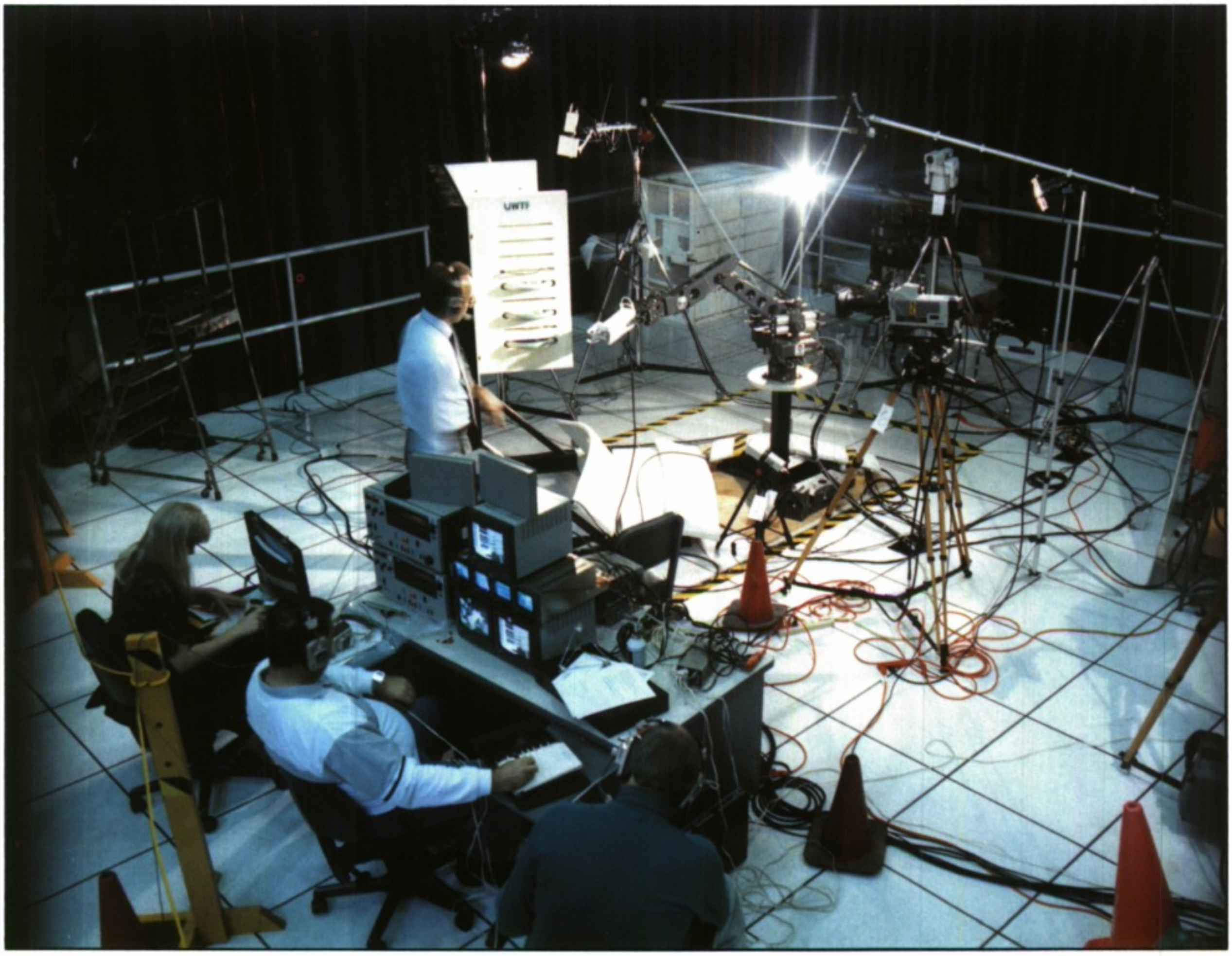
7. Lancaster, F. W. (1968). Information retrieval systems. New York: Wiley.

8. Meister, D., and Mills, R. G. (1971). Development of a human performance reliability data system. In Annals of Reliability and Maintainability, 1971. Tenth Reliability and Maintainability Conference, Anaheim, CA. June 27-30, 1971, 425-439.

9. Melton, A. W., and Briggs, G. E. (1960). Engineering psychology. In Annual Review of Psychology, 11, 71-78.

10. Woodson, W. E. (1981). Human factors design handbook. New York: McGraw-Hill Book Company.

Remote Operator Interaction



In the Remote Operator Interaction Lab, researchers design and conduct experiments and evaluations dealing with human informational needs during the use of telerobots and other remotely operated systems. Operators use various hand controllers with television and direct visual feedback to perform remote manipulation tasks.

HAND CONTROLLER COMMONALITY EVALUATION PROCESS

Mark A. Stuart, John M. Bierschwale, Robert P. Wilmington,
Susan C. Adam, and Manuel F. Diaz
Lockheed Engineering and Sciences Company

Dean G. Jensen
NASA Lyndon B. Johnson Space Center

INTRODUCTION

Hand controller selection for NASA's Orbiter and Space Station *Freedom* is an important area of human-telerobot interface design and evaluation. These input devices will control remotely operated systems that include large crane-like manipulators (e.g., Remote Manipulator System or RMS), smaller, more dexterous manipulators (e.g., Flight Telerobotic Servicer or FTS), and free flyers (e.g., Orbital Maneuvering Vehicle or OMV). Candidate hand controller configurations for these systems vary in many ways: shape, size, number of degrees-of-freedom (DOF), operating modes, provision of force reflection, range of movement, and "naturalness" of use. Unresolved design implementation issues remain, including such topics as how the current Orbiter RMS rotational and translational rate hand controllers compare with the proposed Space Station *Freedom* hand controllers, the advantages that position hand controllers offer for these applications, and whether separate hand controller configurations are required for each application.

Common Space Station and Orbiter hand controllers are desirable for many practical reasons. Common hand controllers would reduce the negative transfer that could occur if many different types of hand controllers were used. The hand controllers need to be selected to minimize astronaut training requirements. Other considerations include the number of controllers required if each system had unique controllers and the associated weight and volume required to accommodate multiple sets and spares.

Several previous studies have evaluated operator performance differences caused by using different hand controller configurations during remote manipulation tasks. For example, O'Hara (1987) compared bilateral force-reflecting replica master controllers to proportional rate six degrees-of-freedom (DOF) controllers during dual-armed remote manipulation tasks and discovered several differences. The six-DOF rate controllers were rated significantly higher in cognitive workload and manual-control workload (ability to control the end effector and the equipment) during dual-armed tasks. O'Hara also reported that the force-reflecting master controller was rated significantly higher in physical workload compared to the six-DOF rate controller. In conclusion, O'Hara found that master controllers resulted in lower performance times and allowed more "natural" control, while six-DOF rate controllers were lower in physical workload. This study was significant, yet limited because only two hand controller types were evaluated under limited operating conditions.

Another relevant study conducted by Honeywell (1989) described current hand controller concepts, the hand controller configurations proposed for Space Station *Freedom*, and the requirements of the space station systems that will use hand controllers. Much of the report was based upon a survey administered to industrial participants, NASA, and universities. A third study (Stuart, Smith, Bierschwale, and Jones, 1989) evaluated the anthropometric and biomechanical interface between test subjects and three and six-DOF joystick and mini-master hand controllers and found that subjects can experience various

types of muscle discomfort due to certain hand controller features. Since these two reports contain little empirical hand controller task performance data, a controlled study is needed that tests Space Station *Freedom* candidate hand controllers during representative tasks. This study also needs to include anthropometric and biomechanical considerations.

EVALUATION

The NASA hand controller commonality evaluation objective was to recommend the hand controller configuration(s) that can meet the Space Station requirements while accomplishing optimal control of each particular system. The recommended configuration(s) shall be chosen to maximize performance, minimize training, and minimize cost of providing safe and productive controls for the Space Station *Freedom* crew.

The hand controller commonality evaluation was conducted as three separate experiments. Experiment One was a non-astronaut hand controller evaluation at three test facilities. Experiment Two was an astronaut hand controller evaluation at the same three test facilities. Experiment Three was a hand controller volumetric evaluation done primarily in the Orbiter and Space Station mockups. All of the evaluations took place at NASA Johnson Space Center (JSC).

EXPERIMENT ONE METHODS

Experiment One was conducted as a repeated measures evaluation (within-subjects design) for each of the six tasks evaluated. These tasks are described in the Apparatus section below. Test subjects used all of the hand controllers for their respective tasks in those modes supported by the hand controllers and the facilities.

SUBJECTS

Twenty-four non-astronaut test subjects were used in Experiment One. Test subjects were partitioned into six independent groups of four test subjects. Each test subject group performed one of the six remote manipulation tasks. Twelve test subjects who had prior dexterous manipulator experience formed three groups, eight test subjects who had prior RMS simulation experience formed two groups, and four test subjects who had prior free flyer experience formed one group.

APPARATUS

Physical Simulations. Physical simulations were performed in the Remote Operator Interaction Laboratory (ROIL). These consisted of the following tasks: fluid quick-disconnect coupling; simulated ORU change-out; and thermal insulation blanket removal. These tasks were performed using a Kraft manipulator slave arm with a JR3 force-torque sensor.

Computer Simulations. Computer simulations took place at two different test sites — the Systems Engineering Simulator (SES) and the Displays and Controls Laboratory (D&CL). The SES tasks were used to investigate rate control mode hand controller characteristics while controlling dynamic free flyer and Space Station Remote Manipulator System (SSRMS) simulations. The specific tasks were OMV docking and logistics module transfer.

The D&CL tasks were used to investigate hand controller characteristics during rate mode for a crane-type manipulator and both rate and position modes for a dexterous manipulator (both kinematic simulations). The D&CL task consists of a sequential SSRMS/dexterous manipulator operation (SSRMS used as a transport device) to perform an ORU replacement task.

Hand Controllers Evaluated. Hand controllers evaluated in this study were provided by NASDA of Japan, McDonnell Douglas/ Honeywell, the Canadian Space Agency, and Goddard Space Flight Center. These hand controllers are illustrated and described in Figure 1.

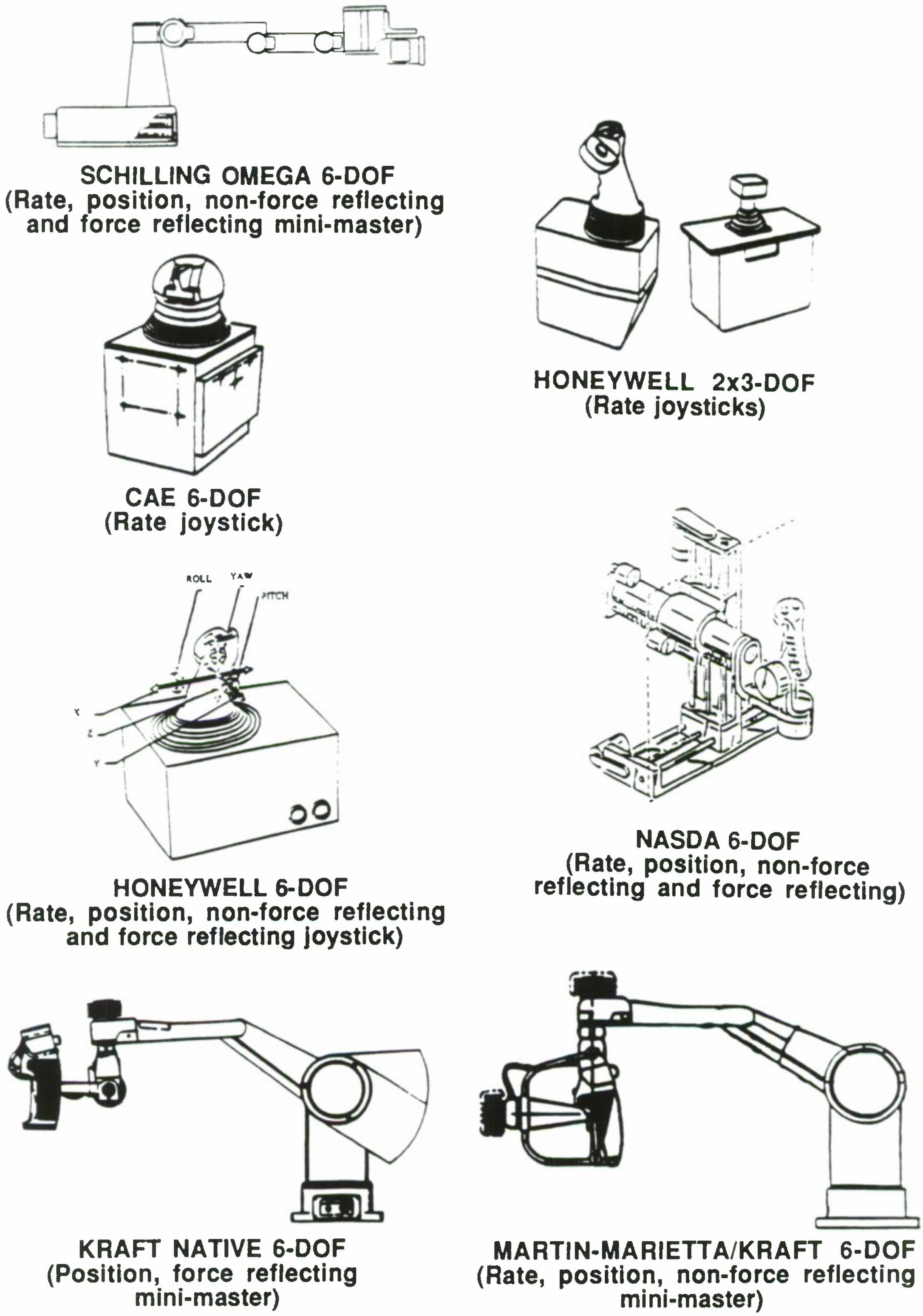


Figure 1. Illustrations and characteristics of hand controllers evaluated.

PROCEDURE

The procedure at each test site included pilot testing with operations-experienced test subjects. At each of the three test sites, test subjects were allowed to switch between different camera views as well as use fine-adjustment camera controls such as focus, pan, tilt, and zoom. The switching and adjusting was done by the test administrator. Tasks at all test sites were broken into their respective subtasks for performance analysis purposes.

ROIL. Test subjects within each of the three dexterous manipulator-experienced groups performed one of the three physically simulated tasks. The ROIL tested position mode with no force reflection (haptic-proprioceptic), position mode with force reflection, and rate mode while operators used a dexterous manipulator. The test subjects followed a prescribed procedure during the performance of the three physical simulation tasks. The subjects used predetermined camera positions of the remote worksite. One of the cameras provided a global view of the entire taskboard area. Camera positions were optimized per task prior to data collection. Test subjects received an equal amount of laboratory training with each of the hand controllers before data collection began. After receiving training for a specific hand controller, each test subject performed the task two times with that controller. The procedure continued in this fashion until subjects within each group performed their respective task twice while using each of the hand controllers. Hand controller use was counterbalanced to control for order effects. After completing two task trials using each hand controller, each subject was administered questionnaires to collect subjective data.

SES. Test subjects within one of the RMS-simulation-experienced groups performed the logistics module transfer task and test subjects within the free-flyer-experienced group performed the OMV task. The SES tested the controllers in rate mode while test subjects used the SSRMS or the OMV. The OMV was controlled in pulse mode and the SSRMS tasks were controlled using the standard proportional rate mode. The subjects used predetermined simulated camera views of the remote worksite as well as a simulated direct view. Subjects completed a familiarization session prior to data collection in the SES and also performed the simulated task two times with each hand controller. Questionnaires were administered after performance of the tasks.

D&CL. Test subjects within the second RMS-simulation-experienced group performed the dual SSRMS/FTS ORU task. The D&CL tested rate and position (non-force-reflective) while operators used a dexterous manipulator in conjunction with the SSRMS. The SSRMS was controlled in the rate mode and the dexterous manipulator was controlled in both rate and in position mode. The subjects used simulated camera views of the remote worksite. After performing the simulated task two times questionnaires were administered.

DATA COLLECTION

Task performance data included the following: time to complete each subtask, reach limits, active hand controller time, the number of hand controller inputs, and error or accuracy counts. Questionnaires were administered to collect the following types of subjective impressions: general acceptability, mental and physical fatigue, and hand controller suitability for specific tasks.

EXPERIMENT TWO METHODS

Experiment Two used astronaut test subjects who performed each of the six tasks at all three test sites.

SUBJECTS

Six crewmembers were used as test subjects in this phase of the evaluation. Prior hand controller experience of each crewmember was assessed.

PROCEDURE

Familiarization with the tasks was required before the crew evaluation took place. This

varied according to the experience level of each individual crewmember. For example, somewhat more familiarization time was necessary for those crewmembers who had no prior OMV or dexterous manipulator task experience.

Each crewmember performed a structured subset of each of the six tasks described in the Experiment One Methods Section. During the task, performance data, such as speed and accuracy, were collected from each crewmember. After performing the structured subset of each of the six tasks with each hand controller, the crewmember was given a brief questionnaire to fill out.

EXPERIMENT THREE METHODS

Experiment Three was a volumetric evaluation which involved astronaut test subjects using all of the hand controllers.

TEST SUBJECTS

Four astronauts performed the evaluations. Attempts were made to have test subjects that range in body sizes from the 95th percentile male to the 5th percentile female (workstations are required to accommodate this range of users).

APPARATUS

Hand controller volumetric evaluations were performed in the Space Station, Cupola, and Shuttle mockups located at NASA JSC. Hand controllers evaluated in Experiment Three are listed in the Experiment One Apparatus section.

PROCEDURE

Single and dual hand controller usage for one operator was addressed at the command and control workstation and the cupola workstation. Side-by-side operator operation was addressed in the cupola. Hand controller mounting and adjustment in the Space Station and Cupola mockups were achieved using two tripods.

DATA MEASUREMENT

Data were collected with both a video recorder and a 35mm camera. Hand controller locations for the various subjects were also recorded. The evaluations consisted of questionnaire administration and anthropometric data collection that addressed the following issues: hand controller swept volume; operator/workstation placement (e.g., crew movement ability in the area); display viewing characteristics (e.g., line of sight characteristics, display obstruction from hand controllers); and reach envelope characteristics (e.g., ability to reach workstation controls). The anthropometric data were incorporated into an analysis of each hand controller configuration within the appropriate workstations on Space Station.

RESULTS

Results of data analyses are summarized as follows: no appreciable astronaut/non-astronaut differences on the performance and subjective data collected; subjective data supported objective (performance) data; trends were consistent across all three tasks conducted; rate control-mode was consistently superior to position control-mode; no advantage demonstrated for force reflection; joystick controllers were superior to mini-master controllers; and the 2x3 DOFs, CAE, and the Honeywell rate-mode were consistently the top hand controller configurations. As a result of these evaluations, a 2x3-DOF hand controller configuration was decreed the Space Station *Freedom* baseline configuration.

REFERENCES

1. Honeywell, Inc. (1989). *Hand controller commonality study -- 289-16007*. Clearwater, Florida.

2. O'Hara, J. (1987). *Telerobotic control of a dextrous manipulator using master and six-DOF hand controllers for space assembly and servicing tasks*. Proceedings of the Human Factors

Society 31st Annual Meeting. Santa Monica, CA: Human Factors Society.

3. Stuart, M. A., Jones, S. F., Smith, R. L., and Bierschwale, J. M. *Preliminary hand controller evaluation*. Houston, TX: NASA Lyndon B. Johnson Space Center, 1989; NASA Tech. Report JSC-23892.

PROGRAMMABLE DISPLAY PUSHBUTTONS ON THE SPACE STATION'S TELEROBOT CONTROL PANEL

Mark A. Stuart, Randy L. Smith, and Ervette P. Moore
Lockheed Engineering and Sciences Company

Research directed by Jay Legendre, Manager, Remote Operator Interaction Lab, NASA JSC.

INTRODUCTION

The Man-Systems Telerobotics Laboratory at NASA's Johnson Space Center and supported by Lockheed, is working to ensure that the Flight Telerobotic Servicer (FTS) to be used on the Space Shuttle (Orbiter) and the Space Station has a well designed user interface from a Human Factors perspective. The FTS, which is a project led by NASA's Goddard Space Flight Center, will be a telerobot used for Space Station construction, maintenance, and satellite repair. It will be directly controlled from workstations on the Orbiter and the Space Station and monitored from a ground workstation. The FTS will eventually evolve into a more autonomous system, but in the short-term the system will be manually operated (teleoperated) for many tasks. This emphasizes the importance of the human/ telerobot interface on this system.

The information driving the design of the FTS control panel is being provided by task analyses, workstation evaluations, and astronaut/FTS function allocations. Due to space constraints on the Orbiter and the Space Station, an overriding objective of the design of the FTS workstation is that it take up as little panel space as possible.

This phase of the FTS workstation evaluation covers a preliminary study of programmable display pushbuttons (PDPs). The PDP is constructed of a matrix of directly addressable electroluminescent (EL) pixels which can be used to form dot-matrix characters. PDPs can be used to display more than one message and to control more than one function. Since the PDPs have these features, then a single PDP may possibly replace the use of many single-function pushbuttons, rotary switches, and toggle switches, thus using less panel space. It is of interest to determine if PDPs can be used to adequately perform complex hierarchically structured task sequences.

Other investigators have reported on the feasibility of using PDPs in systems design (Hawkins, Reising, and Woodson, 1984; and Burns and Warren, 1985), but the present endeavor was deemed necessary so that a clearly defined set of guidelines concerning the advantages and disadvantages of PDP use in the FTS workstation could be established. This would ensure that PDP use was optimized in the FTS workstation.

The objective of this investigation was to compare the performance of experienced and inexperienced Remote Manipulator System (RMS) operators while performing an RMS-like task on simulated PDP and non-PDP computer prototypes so that guidelines governing the use of programmable display pushbuttons on the FTS workstation could be created. The functionality of the RMS on the Orbiter was used as a model for this evaluation since the functionality of the FTS at the time of this writing has not been solidified.

METHOD

APPARATUS

Computer prototyping was used as the means of evaluating the two different FTS control panel layouts. Hypercard was used as the prototyping package and it was run on an Apple Macintosh computer. Hypercard was also used as a data acquisition package once testing began. Total task time and the total number of commands activated were recorded.

The simulated task consisted of the operations to deploy a satellite on the Space Shuttle. This task required simulated RMS joint mani-

pulations, camera manipulations, as well as other RMS-like activities, while using the computer prototypes.

The non-PDP control panel is depicted in Figure 1. The distinguishing feature of this configuration is that traditional single-function pushbuttons are used in conjunction with a simulated EL panel to activate commands. The EL panel was simulated in this evaluation by displaying single-function commands as they would appear on the EL panel in the upper right-hand corner of the prototyped screen. The simulated EL panel was used because the space constraints of the Macintosh computer would not allow the display of all of the functionality at one time. This then made it possible to study a task as complex as an RMS-like operation on this particular microcomputer.

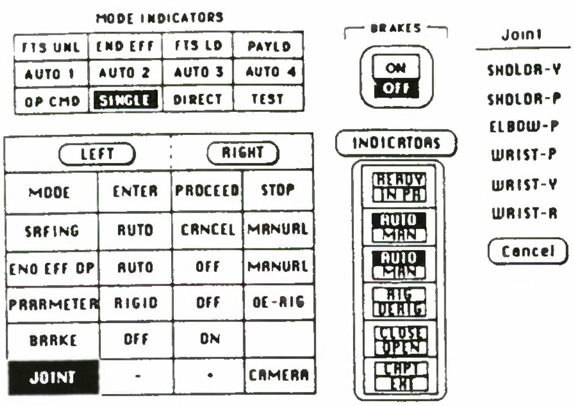


Figure 1. Non-PDP control panel prototype.

The PDP control panel is depicted in Figure 2. This control panel utilized simulated PDPs instead of single-function pushbuttons. In Figure 2, the PDPs are the twelve pushbuttons located in the lower-middle portion of the display. The portions to the left and top of the display are status indicators that were used to display various functional states.

When a PDP is selected, the name of that function is then displayed in a small simulated EL display located just above the PDP cluster and the options that follow within that functional category are then displayed by the PDPs. For example, when SINGLE is selected in Figure 2, the display changes to that depicted in Figure 3. In Figure 3, SINGLE is now displayed in the EL display and the PDPs have changed to list the options that follow under SINGLE. The small EL display was designed to serve as a navigational aid to help orient operators throughout performance of the hierarchically structured tasks. It was contended that the use of the navigational aid in the PDP hierarchy would be useful since a previous evaluation (Gray, 1986) found that navigational aids are helpful with hierarchical search tasks through menu structures on a computer.

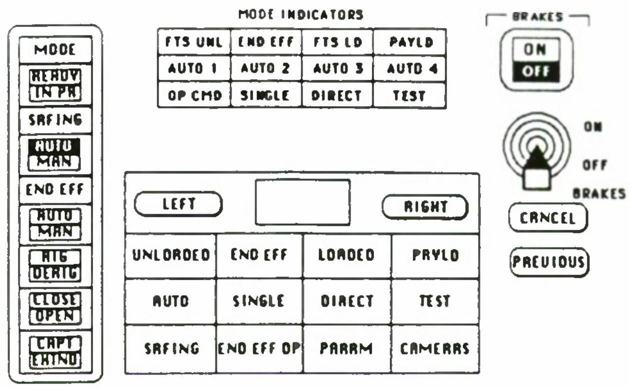


Figure 2. PDP control panel prototype.

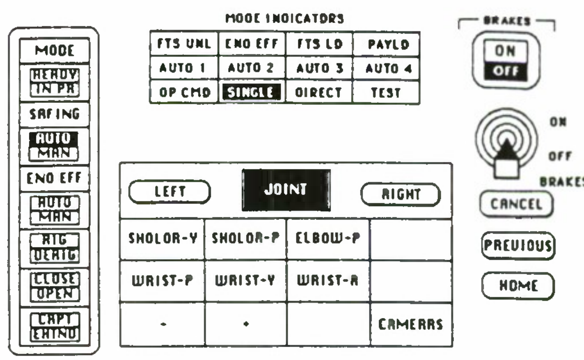


Figure 3. PDP control panel with PDP changes and navigational aid.

It was determined that there was a problem in maneuvering across functional modalities during the development of the PDP prototypes because it required many commands to do so. For example, when one was in the RMS joint manipulation mode, it would require several steps, including going back to the "Home" level of the task hierarchy first, to be able to make camera adjustments. Since the RMS operation requires much maneuvering across modalities during its use, this PDP arrangement would result in many circuitous movements and much wasted time. Therefore, a special PDP was developed for the present investigation which would readily allow operators to "jump" across functional modalities with a single command located within the PDP matrix.

EXPERIMENTAL VARIABLES

The independent variables in this investigation were the two different RMS operating experience levels of the subjects (experienced and novice) and the two different control panel prototypes (non-PDP and PDP). Since each subject was tested on each of the two control panel prototypes (in counterbalanced order), then a 2 x 2 repeated measures experimental design was used. Of specific interest was the comparison between PDP and non-PDP usage, the difference in the performance and subjective impressions of the two different subject groups, the use of navigational aids, and informational needs of operators while performing simulated FTS tasks.

The dependent variables were operator task completion times, number of commands required to complete the task for each control panel prototype, a question of preference between the two different control panel prototypes, questionnaire responses, and subjective impressions.

SUBJECTS

Volunteer subjects from both Johnson Space Center and Lockheed took part in this experiment. Four subjects who had no prior RMS training comprised the novice users group. Four subjects who had completed training on a high-fidelity simulator of the RMS comprised the experienced users group.

PROCEDURE

Performance of a simulated RMS-like task scenario was used for each of the control panel configurations. Each scenario covered simulated RMS-like manipulation activities and the testing took place on the Apple Macintosh SE computer. The task scenario was identical for both control panel configurations.

Before testing began, each subject had the basic functionality of each of the control panels explained to them. Subjects then completed a practice session on the first control panel configuration that they would be using. A subject would then perform the simulated tasks on the Macintosh. After the subject's first scenario was completed, the same procedure was followed using the other control panel prototype. Order effects were controlled by having an equal number of subjects begin the testing with the non-PDP control panel as those who began the testing with the PDP control panel within each of the two subject groups.

After performing the task scenarios on both of the control panel prototypes, each subject was asked to select which of the two control panel prototypes were preferred. Each subject was also asked to complete a questionnaire designed to garner subjective impressions concerning the control panels. Subjects rated five issues on a five-point Likert-scale where one point indicated "Least" and five points indicated "Most." These five issues were "Maneuver Across Modalities," "Maneuver Within Modalities," "Provides Task Structure," "Contributes to Task Structure," and "Ability to Make Commands." Subjects then answered open-ended questions concerning PDP use.

RESULTS AND DISCUSSION

Data were collected and analyzed with the objective of determining differences in user performance and preferences between the two different control panel configurations so that,

ultimately, guidelines concerning the use of PDPs could be established. All numeric data were statistically analyzed with a repeated measures analysis of variance.

Analysis of the performance data revealed that subjects used significantly ($p = 0.001$) fewer commands when using the PDP control panel prototype than they did while using the non-PDP control panel. Interestingly, though, subjects did not significantly differ in the amount of time that it took for them to complete the two tasks. The average task time for the PDP prototype was 18:12 while it was 18:49 for the non-PDP condition. This finding provides some support for the PDP prototype in the sense that if more commands are required to perform the same task in virtually the same time frame then the condition which requires more commands to be activated may predispose operators to make more errors.

Analysis of the subjects' control panel preferences revealed that all eight of the subjects preferred the PDP control panel over the non-PDP control panel. As Table 1 indicates, the analysis of the five-point Likert-scale questionnaire responses also provided strong support for the PDP control panel since subjects rated two of the five questionnaire items significantly ($p < 0.05$) higher for the PDP prototypes. These two questionnaire items were "Maneuver Within Modalities" and "Ability to Make Commands." Subjects also rated the PDP prototype higher on the other three questionnaire items, although these differences were not statistically significant. There was also statistical significance ($p = 0.049$) due to the RMS experience level of the subjects where the novice users had a higher rating on the "Maneuver Across Modalities" question.

Subjective comments were also collected from each of the subjects. These are summarized in Table 2. The comments were categorized as either positive or negative with respect to PDP usage.

The subjective impressions indicate that PDPs can have very good as well as very bad features. It was observed by the subjects that the use of the PDPs can result in less panel space used and that they can provide task structure in the sense that they can clearly delineate what task options are available at specific times. On the negative side, subjects expressed that one loses "global perspective" with the use of PDPs and that this can contribute to task disorientation. It was also stated that PDPs should not be used in "exceedingly" complex systems.

Subjective impressions were also studied to determine if there was a difference between the two RMS-experience groups. Data analysis revealed that there were no differences since the comments were common across both groups.

TABLE 1.

Five-point Likert-scale responses for the non-PDP and PDP control panel prototypes

Questionnaire Item	Control Panel Non-PDP	PDP
Maneuver across modalities	3.12	3.87
Maneuver within modalities	2.87	4.25 *
Provides task structure	2.75	4.12
Contributes to task orientation	2.62	3.50
Ability to make commands	3.00	3.87 *

* Significant at $p < 0.05$

CONCLUSIONS

The ultimate objective of this investigation was to establish a set of guidelines concerning the use of PDPs for the FTS workstation. The

data collected during this investigation were then used to create this set of guidelines. It is contended that the established set of guidelines will also be generalizable to other workstations as well. These guidelines are listed in Table 3.

It is clear from the previously mentioned experiment results and subjective comments that the use of PDPs does in fact present a trade-off — there is some good as well as some bad about them. It is for this reason then that PDPs should be used judiciously because improper usage can contribute to task complexity and user task-disorientation. It is contended that the previously mentioned set of guidelines will help to ensure that PDPs will be optimally designed and arranged.

TABLE 2.

Positive and negative subjective impressions concerning PDP usage

---

Positive
- Provide task structure
- Save panel space
- User attention is more localized
- Good when working within a functional modality (e. g., camera manipulation)
- Navigational aids provide user guidance
- Good for infrequently used sub-tasks
- Can result in reduced search time

Negative
- Processing time (option refresh rate) to perform next steps was too slow
- Bad if used in highly complex systems (e. g., large number of functional modalities within the overall task)
- Lose global perspective because fewer spatially redundant cues
- Not good for applications where few controls are used frequently
- Possibility of getting lost in complex task structures
- May result in more cognitive processing

Future research endeavors should examine the use of actual, hard-wired PDPs in full-scale mockups while performing high-fidelity simulated tasks. This would increase the external generalizability of the results. The development of an equation which would precisely determine how many PDPs should be used for a specific task may be possible. This equation would have to take into account variables such as the frequency that all of the commands are activated, as well as the depth and breadth of the task hierarchical structure.

TABLE 3.

Guidelines concerning PDP usage

- Use PDPs instead of other controls if PDP usage reduces the total number of commands to perform the task and doesn't significantly increase task completion time.
- A PDP or control capability should be provided that will allow "jumping" across functional modalities
- Navigational aids should be used to help orient users
- May be better for infrequently used sub-tasks
- May be better when working within a functional modality
- Should not be used for certain critical functions, such as brake control
- Should give an indication of the number of hierarchical steps the operator is away from the "Home" level

---

ACKNOWLEDGEMENTS

Support for this investigation was provided by the National Aeronautics and Space Administration through Contract NAS9-17900 to Lockheed Engineering and Sciences Company.

REFERENCES

1. Burns, M. J., and Warren, D. L. (1985). Applying programmable display pushbuttons to manned space operations. In *Proceedings of the Human Factors Society 29th Annual*

Meeting (pp. 839-842). Santa Monica, CA: Human Factors Society.

2. Gray, J. (1986). The role of menu titles as a navigational aid in hierarchical menus. *SIGCHI Bulletin, 17(3),* 33-40.

3. Hawkins, J. S., Reising, J. M., and Woodson, B. K. (1984). A study of programmable switch symbology. In *Proceedings of the Human Factors Society 28th Annual Meeting* (pp. 118-122). Santa Monica, CA: Human Factors Society.

4. Stuart, M. A., Smith, R. L., and Moore, E. P. (1988). *PDP preliminary evaluation* (Lockheed EMSCO Memo No. 88-1078). Houston, TX: Lockheed Engineering and Management Services Company.

SPEECH VERSUS MANUAL CONTROL OF CAMERA FUNCTIONS DURING A TELEROBOTIC TASK

John M. Bierschwale, Carlos E. Sampaio, Mark A. Stuart, and Randy L. Smith
Lockheed Engineering and Sciences Company

Research directed by Jay Legendre, Manager, Remote Operator Interaction Lab, NASA JSC.

INTRODUCTION

Telerobotic workstations will play a major role in the assembly of Space Station *Freedom* and later in construction and maintenance in space. Successful completion of these activities will require consideration of many different activities integral to effective operation: operating the manipulator, controlling remote lighting, camera selection and operation, image processing, as well as monitoring system information on all of these activities.

Of these activities, the vision (camera viewing) system is particularly important. During many tasks where a direct view is not possible, cameras will be the user's only form of visual feedback. If the vision system is manually controlled and both hands are busy during the performance of a dynamic task, it will require reorientation of the hands and eyes between the manipulator controls, vision system controls, and view of the remote worksite. Allocating some or all of the control of vision system components to voice input may lessen the workload of the operator, reduce movement time, and ultimately improve performance. Voice input is currently being considered for this as well as other applications by NASA.

Very few studies are found in the literature that investigate the use of voice input for camera control. The only study that was found (Bejczy, Dotson, Brown, and Lewis, 1982) was relevant in that it investigated voice input for camera control of the Remote Manipulator System (RMS) and payload bay cameras used on the Space Shuttle. Although statistical analyses were not presented, voice input was found to be 10% slower across four subjects.

The philosophy of the present investigation differs in that subjects were not constrained to current RMS control panel terminology and organization. Subjects used words from a vocabulary sheet developed in a previous study (Bierschwale, Sampaio, Stuart, and Smith, 1989) to construct camera commands to accomplish a telerobotic task. The subjects' vocabulary preferences are presented elsewhere (Bierschwale, et al., 1989).

It is important to consider current terminology so that personnel are not forced to learn new jargon. However, the use of voice input was not considered in the development and selection of the current terminology and switch labels. Choice of vocabulary is very important in terms of recognizer performance and user acceptance. Successful vocabulary design (ultimately the human machine interface design) will most readily be achieved by considering the recognition qualities of the commands and cognitive relationship between the commands and their respective actions.

A potential problem with voice control of cameras may be verbalizing the directions to move the cameras. Many people have difficulty when providing verbal directions. An example would be saying "left" when "right" is meant. Indeed, this cognitive difficulty when verbalizing directions has been noted with voice control of cursor movement while editing text (Murray, Praag, and Gilfoil, 1983; and Bierschwale, 1987).

Identification of critical issues such as this early in the design phase will allow for more effective implementation of a voice

commanded camera control system. In more general terms, one report (Simpson, McCauley, Roland, Ruth, and Williges, 1985, p. 120) found that, historically, "projects designed from inception to incorporate a voice interactive system had a greater probability of success than when the capability was added to an existing system." By understanding the differences between the two modes of input, a more effective utilization can be made of both voice and manual input.

The objectives of this study are as follows: (1) optimize the vocabulary used in a voice input system from a Human Factors perspective, (2) perform a comparison between voice and manual input in terms of various performance parameters, and (3) identify factors that differ between voice and manual control of camera functions.

METHOD

SUBJECTS

Eight volunteer subjects were selected to participate in this evaluation. These subjects were partitioned into the following two groups: an experienced group of four subjects who were familiar with telerobotic tasks and workstations and an inexperienced group of four who were not familiar with these concepts.

APPARATUS

Testing took place in the Man-Systems Telerobotic Laboratory (MSTL) located at the NASA Johnson Space Center. A Kraft manipulator slaved to a replica master controller was used to perform a remote telerobotic task. The task selected for this study was a generic pick and placement task. This task required a high degree of visual inspection and dextrous manipulation. The tasksite is depicted in Figure 1.

Two 4-inch tall and two 10-inch tall tiers were placed on a semicircular taskboard in front of the Kraft manipulator. Three task pieces were placed on the lower left-hand tier and three were placed on the upper left-hand tier. On the right-hand side, four receptacles were placed on the upper and lower tiers (two receptacles per tier). The task consisted of locating, grasping, transporting, and depositing each of four task pieces into the correct receptacle. In addition to the required manipulation, subjects had to move cameras, adjust lens parameters, and select views to successfully complete the task. During the task, subjects were instructed which task piece and receptacle were involved.

Two cameras equipped with remote pan, tilt, zoom, focus, and iris controls provided the operator with two oblique views of the worksite (i.e., approximately 45 degrees above the horizontal plane with one displaced 45 degrees to the left and the other camera 45 degrees to the right). A fixed-focus camera provided a "bird's-eye" view of the entire work area looking down at a 45-degree angle on the worksite from above the task. The two oblique views were input to a 21-inch monitor where only one view could be shown at a time. The "bird's-eye" view was continuously displayed on a 9-inch monitor positioned atop the larger monitor.

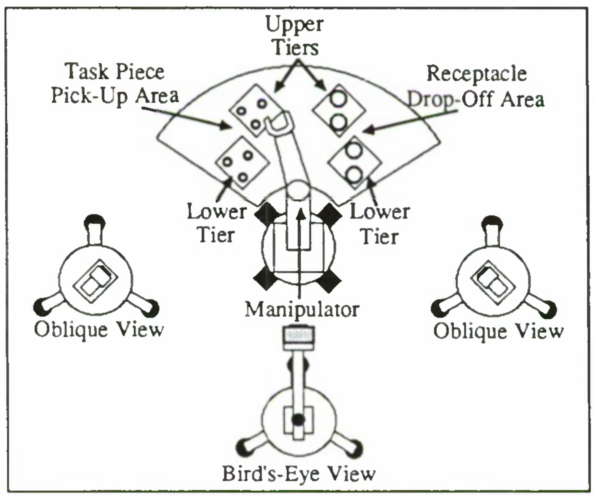


Figure 1. Overhead view of remote work site.

The left oblique view showed the task pieces and the surrounding area. The right oblique view showed the box and the surrounding area. The "bird's-eye" view showed the entire work area. Taskpieces were aligned such that the left oblique view was required to read their

markings while the right oblique view was required to read the receptacles' markings.

A practice task (using direct view) was devised so subjects could become familiar with the manipulator controls, camera views, and kinesthetics of the arm movements and positions that they would be using later during data collection. The practice task used the identical views, taskboard, tier placement, and similar task objectives.

During use of manual input, the camera controls were placed directly in front of the subjects. Subjects were required to use their right hand to operate both the manipulator and camera controls. This required halting the manipulator to operate the cameras. This was a simulation of a hands-busy scenario where voice input might aid performance.

A vocabulary list containing stereotypical words determined in a previous evaluation (Bierschwale et al., 1989) was used for voice input. A separate vocabulary sheet was used for each subject and the words were randomly listed under each icon (descriptive of the camera function) to avoid any possible list order effect.

In order for the control system to be flexible enough to accommodate the various word combinations, an experimental approach was used that has been referred to as the "Wizard of Oz" method. This is often used in user-computer interaction research and is summarized in Green and Wei-Haas (1985). For this evaluation, a "wizard" carried out the actions of a speech recognizer. This method has been used before with voice input research. One study (Casali, Dryden, and Williges, 1988) used a wizard recognizer to evaluate the effects of recognition accuracy and vocabulary size on performance.

The "wizard" was situated at the camera controls out of the field-of-view behind and to the right of the subject. When voice input was used, the "wizard" wore a headset which allowed him to screen out external noise and concentrate on the commands issued by the subject through a microphone.

VARIABLES

Three independent variables with two levels each were studied in this evaluation: input modality (voice or manual), level of experience (experienced or inexperienced), and administration order (voice followed by manual input or manual followed by voice input). Experience level and administration order were between-subjects variables while input modality was a within-subjects variable.

Dependent variables consisted of task completion time, number of camera commands, and errors. Scaled question and questionnaire responses were also collected.

PROCEDURES

At the beginning of the evaluation, subjects were provided with a brief explanation of the purpose of the study. Each subject received instruction on the use of the manipulator controls that would be needed to perform the manipulation tasks.

Following performance of the practice task (using direct view), a videotape was used to illustrate the different camera and lens movements that would be available on the two adjustable cameras. The investigator used deliberate wording when pointing to the corresponding icons on either the camera control panel (manual) or the vocabulary sheet (voice), so as not to bias any subject's selection of vocal commands. Prior to each of the two conditions, subjects were instructed on the use of the respective camera controls. When using manual input, a template with descriptive icons illustrating the functions was placed over the controls so that the subjects would not be biased in their vocabulary selection (for voice input) by using the listed labels. These same icons were used on the vocabulary sheet for voice input.

The subjects' view of the task was then obstructed so that they had to rely totally on the camera views. Each subject performed two sessions under both conditions (voice and manual input). The first was a practice session using an abbreviated version of the task with

the second being the complete task. This practice also allowed subjects, while using voice input, to become familiar with the designated words and select the few they might prefer to use. Administration order was counterbalanced with half of the subjects using voice input first and half using manual input first. Additionally, to avoid any memorization of task requirements, different locations and task piece selections were used for each of the four sessions (one practice and data collection session per condition).

Following the practice session, the data collection session was conducted. Subjects were instructed to work quickly while making as few errors as possible. If an error occurred, the taskpieces and receptacles were placed in the configuration present prior to the error and the subject repeated the trial. While setup time was not recorded, repetition of the trial was included in the completion time. The video images used to perform the task (excluding the "bird's-eye" view) were recorded along with audio input from the subjects' headset.

Following completion of the data collection session for each condition, subjects completed a questionnaire. The procedure for the second condition progressed in the same manner. After testing was finished, another questionnaire, involving comparison of the two modalities, was completed by the subjects.

RESULTS AND DISCUSSION

PERFORMANCE DATA

Analysis of Variance (ANOVA) results are presented for the task completion times, number of commands, and errors. Table 1 presents the group means for each of these measures.

An ANOVA run on the task completion times found that voice input was significantly slower than manual input for controlling cameras in this task ($F(1,4) = 19.80$, $p < .05$).

In order to allow for a direct comparison of the number of camera manipulations, the voice commands were tallied such that a single command consisted of both activating and stopping the movement (actually two voice commands issued). An ANOVA run on the number of commands that were used found that significantly more commands were used with manual than voice input ($F(1,4) = 10.34$, $p < .05$).

It was expected that more manual commands would be used since people tend to "bump" manual controls and set things up perfectly. With voice input, subjects tended to accept coarse adjustments because of the difficulties imposed by the system lag time and lack of variable rate control. If examined in conjunction with the task completion times, it is seen that subjects used more time to execute fewer commands with voice input. It may very well be that using voice input to control the cameras resulted in more cognitive difficulty associated with each command which could result in more errors. On the other hand, assuming a constant error rate, the greater number of commands given with manual input would increase the probability that an error will occur. If this effect exists, this is a system trade-off that will need to be evaluated.

TABLE 1.

Group means for performance measures.

	Voice input		Manual input	
	Exp.	Inexp.	Exp.	Inexp.
Completion Time (Minutes)	12.58	15.46	10.94	12.41
Commands *	90.30	104.50	114.00	130.30
Manipulation Errors	.75	.75	1.00	1.00
Focusing Error Rates (percent)	30.50	32.80	50.00	37.00

* Does not include extra commands resulting from directional errors.

It was hypothesized that fewer manipulator errors would occur with voice control since this would allow the subjects to keep their eyes on the screen and avoid interruption of the task. However, the results of an ANOVA

show that there was not a statistically significant difference in the number of manipulation errors between modalities. The makeup of the task was such that few errors were committed with either modality.

A directional error consisted of moving one direction when one wanted to move in the opposite direction. Very few directional errors were observed across the functions except when focusing the cameras. More errors were made when using manual control to focus the cameras than when using voice control. However, results of an ANOVA revealed no significant difference in focusing error rates between the two input modalities.

The probable reason for the high focusing error rates was that the task required zooming the focal length back and forth and the subject would usually guess which directional command would bring the picture into focus. Possible reasons why somewhat higher error rates occurred during manual input were that subjects tended to perform more commands, as was previously mentioned, and were more likely to attempt to bring the picture into exact focus. With voice input, focusing was difficult due to the sensitivity of the focusing operation and the system lag time. Subjects would often accept a less than perfect image.

SUBJECTIVE RESPONSES

The following types of questions were asked concerning the two input modalities: scaled questions, open-ended questions, and yes/no questions that allowed the subject to elaborate. Analysis revealed no real differences in preference between the two modalities of input and two experience groups across all of the questions for this task. However, similar subjective comments, concerning advantages and disadvantages of the two modalities for performance of this task, were frequently made across many of the questions and are summarized in Table 2.

When subjects were asked what telerobotic workstation functions they would recommend allocating to voice input, the following applications were given: selecting or moving cameras, controlling lights, halting the manipulator arm, setting the manipulator grip lock, changing modes, and panning and tilting only. For the most part, these applications are of a discrete nature that minimize the disadvantages of voice input listed in Table 2.

TABLE 2.

Advantages and disadvantages of voice-operated camera control.

VOICE INPUT	
ADVANTAGES	DISADVANTAGES
Hands and eyes free Good for single, gross movements while hands are occupied Possibility for simultaneous camera/ manipulator control	Cognitive difficulty verbalizing commands/directions System lag time Two step start-stop process Can't perform two camera movements at once
MANUAL INPUT	
ADVANTAGES	DISADVANTAGES
Finer positioning than voice input Less mental load than voice input Quicker system response time	Diverting eyes and hands from telerobotic task to adjust camera controls

CONCLUSION

This investigation has evaluated the voice-commanded camera control concept. For this particular task, total voice control of continuous and discrete camera functions was significantly slower than manual control. There was no significant difference between voice and manual input for several types of errors. There was not a clear trend in subjective preference (across several questions) of camera command input modality. Task performance, in terms of both accuracy and speed, was very similar across both levels of experience.

One problem that emerged was that numerous focusing errors (30-50%) were observed across both groups and modalities. For tasks as dynamic as this, development of an

autofocusing system is highly recommended to avoid operator frustration and inefficiency.

The fundamental advantage that voice input had over manual input, as mentioned by both groups of subjects, was that it allowed the hands and eyes to be free to do other tasks.

Unfortunately, voice input of camera controls also resulted in cognitive difficulty when verbally transcribing movements, specifying the correct directions, and stopping movements. The advantage of manual input was that it allowed precise positioning. The applications that subjects suggested for voice input at a telerobotic workstation were of a discrete control nature.

Most of the problems seem to be associated with the movement processes. While each distinct movement (zoom, pan, tilt, etc.) was not directly compared across both modalities, subjective comments indicate that the problem is a fundamental one of verbal control of a spatial motor task. The study by Bejczy, et al. (1982) also stated that controlling camera movement was troublesome for the subjects.

The results of this investigation indicate that using voice input for control of discrete types of camera operations (selecting cameras, multiplexing, and selecting rates) could aid performance in a telerobotic task. Control of continuous camera functions by voice input is not recommended.

A combination of voice input and manual input for control of camera movements would take advantage of the best aspects of each of the control modalities. Future studies should evaluate alternate methods of controlling camera movements. Some examples are: (1) a hand-controller mounted joystick whose function is selected (camera pitch, camera roll, zoom, focus, and iris control) by voice and controlled manually, which would save panel space by only requiring one control for each or all of the cameras, (2) activating movements by voice and stopping them manually using a switch on the hand controller, and (3) use discrete levels of zoom, focus, and iris (Level 1, 2, 3, 4, etc.) and discrete movements of cameras (perhaps angular, as in pan right 30°, 60°, etc.). Other modalities of input such as an eye tracking device or head-slaved camera control device should also be investigated.

These results were achieved with a particular task, manipulator, and camera control system. A voice recognizer simulation was used that had the advantage of 100% recognition and the possible disadvantage of slower response time. An actual voice recognizer will not perform this well. With decreasing recognition rates, several things will probably occur (although it is difficult to precisely quantify the magnitude of the effects). For example, one study (Casali et al., 1988) found a 17% increase in completion time for a data entry task when recognition rate dropped from 99 to 95% and a 50% increase in completion time when the rate dropped from 99 to 91%. It was also found that each lowered level of recognition produced a significant decline in subjective acceptance of the system.

Different tasks and control systems might produce different results, although it is believed that the trends discussed in this report are applicable across a wide variety of telerobotic tasks. Thus, it is contended that the results will have immediate application to the design of the telerobotic workstations.

ACKNOWLEDGEMENTS

Support for this investigation was provided by the National Aeronautics and Space Administration through Contract NAS 9-17900 to Lockheed Engineering and Sciences Company.

REFERENCES

1. Bejczy, A. K., Dotson, R. S., Brown, J. W., and Lewis, J. L. (1982). Voice control of the space shuttle video system. In *Proceedings of the 17th Annual Conference on Manual Control* (pp. 627-640). Pasadena, CA: Jet Propulsion Laboratory and California Institute of Technology.

2. Bierschwale, J. M. (1987). *Speech versus keying commands in a text-editing task.* Unpublished master's thesis, Texas A&M University, College Station, TX.

3. Bierschwale, J. M., Sampaio, C. E., Stuart, M. A., and Smith, R. L. *Development of a vocabulary for voice input of camera controls during a telerobotic task.* (NASA Tech. Report JSC - 23706). Houston, TX: NASA Lyndon B. Johnson Space Center.

4. Casali, S. P., Dryden, R. D., and Williges, B. H. (1988). The effects of recognition accuracy and vocabulary size of a speech recognition system on task performance and user acceptance. In *Proceedings of the Human Factors Society 29th Annual Meeting* (pp. 232-236). Santa Monica, CA: Human Factors Society.

5. Green, P., and Wei-Haas, L. (1985). The rapid development of user interfaces: experience with the Wizard of Oz method. In *Proceedings of the Human Factors Society 29th Annual Meeting* (pp.470-474). Santa Monica, CA: Human Factors Society.

6. Murray, J. T., Van Praag, J., and Gilfoil, D. (1983). Voice versus keyboard control of cursor motion. In *Proceedings of the Human Factors Society 27th Annual Meeting* (pp. 103). Santa Monica, CA: Human Factors Society.

7. Simpson, C. A., McCauley, M. E., Roland, E. F., Ruth, J. C., and Williges, B. H. (1985). System design for speech recognition and generation. *Human Factors*, *27*, 115-141.

SPACE STATION *FREEDOM* COUPLING TASKS: AN EVALUATION OF THEIR TELEROBOTIC AND EVA COMPATIBILITY

Carlos E. Sampaio, John M. Bierschwale, Terence F. Fleming, and Mark A. Stuart
Lockheed Engineering and Sciences Company

Research directed by Jay Legendre, Manager, Remote Operator Interaction Lab, NASA JSC.

INTRODUCTION

After its completion, Space Station *Freedom* will continue to require a great deal of maintenance and support work in order to maintain daily operations. Dextrous manipulators including the Flight Telerobotic Servicer, the Special Purpose Dextrous Manipulator, and the Japanese Experimental Module Fine arm will not only be critical to the performance of these tasks but may actually be the primary system devoted to the execution of many of them.

Among the tasks to be commonly performed will be the coupling and uncoupling of fluid connectors designed to provide remote resupply of liquids and gases in orbit (NASA, 1989). This will be done using various quick disconnect (QD) couplings designed to mate and demate repeatedly without leakage. At present, several designs exist which allow the couplings to be quickly mated and demated by an extravehicular astronaut. While it is critical that these couplings be capable of manipulation by the suited astronaut, it is equally critical that these couplings be capable of successful operation with a telerobotic manipulator in order to reduce the likelihood of these hazardous extravehicular operations in the first place. Consequently, these couplings necessitate a design that is compatible with both modes of operation.

QD coupling designs and methods of actuation can vary widely. The coupling's contents, the amount of pressure it will have to sustain, the amount of flow it will need to accommodate, as well as several other factors all have a bearing on the coupling's final form. Clearly aboard *Freedom*, the varying conditions under which the different QDs operate will necessitate that their designs be different as well. Just as clear, however, is the concern that a proliferance of coupling designs will, at best, often result in uncertainty in a coupling's operation when encountered, and at worst, result in unsuccessful mating or even loss of fluid or pressure as a result of implementing the incorrect coupling process. Although the size and action of the couplings will clearly need to vary, it is preferable that a similar operation concept might be shared over the coupling points aboard *Freedom* in order to reduce the likelihood of using the incorrect procedure.

It is widely held that in the vast majority of cases, a task that has been designed to be telerobotically compatible will be compatible with the extravehicular astronaut as well (Newport, 1989). This study, conducted in the Remote Operator Interaction Laboratory (ROIL) at NASA's Johnson Space Center (JSC), evaluated subjects' abilities to mate and demate QD couplings of varying design both telerobotically as well as manually. In a previous study assessing various telerobotic control modes, a manual condition was included as a representation of the optimal performance to strive for in the design of a space glove (Hannaford, 1989). Therefore, the manual condition in this study is similarly included as a baseline to reasonably approximate extravehicular activity (EVA).

In collaboration with various telerobotic interface development facilities including the ROIL, Symetrics Inc. has been iteratively designing fluid couplings whose operation is intended to be telerobotically as well as EVA compatible. One of these iteratively designed couplings was among the four coupling designs evaluated in this investigation. Thus the hypothesis of this study proposes that the

coupling designed to be telerobotically and EVA compatible will be mated and demated the most quickly and be most preferred subjectively for both the telerobotic as well as manual conditions.

METHOD

SUBJECTS

Four subjects participated voluntarily in this study. In order to minimize learning effects associated with the various systems involved, all subjects had extensive experience with the telerobotic and viewing systems employed in the study. None of the subjects had any experience operating the QD couplings prior to their participation.

APPARATUS

Three equipment systems were employed in the ROIL. These were a telerobotic system, a viewing system, and a task support structure. The telerobotic system consisted of a Kraft force-reflecting master-slave manipulator. The viewing system consisted of three camera views displayed on three monitors, two of which were 21-inch monitors with one 9-inch monitor. The 21-inch monitors displayed close-up views of the couplings from both front and rear, while the 9-inch monitor displayed an overall view providing the subject with information regarding the orientation of the manipulator to the task piece. The task support structure consisted of a 72-inch by 48-inch metal frame upon which each coupling was attached one at a time during testing. As demonstrated in Figure 1, the designs of the four couplings included in this study differed quite a bit, as did their actuation.

Coupling A was demated by grasping the outer sleeve between the two flanges and applying axial force toward the flex hose. It was demated when enough force was applied to overcome the breakout force of the coupling. Mating occurred by aligning the coupling onto the nipple end and applying axial force until the outer sleeve locked back into place. This was the customized coupling designed specifically by Symetrics to be telerobotically and EVA compatible. The flanges of the outer spool-shaped sleeve were designed to be slightly wider than the telerobotic grippers. This allowed some compliance in grappling the fixture while still providing a sufficient brace in order to apply the axial force necessary for demating and mating. Another aspect of coupling A's design which did not exist on the other couplings was a chamfering of the entrance at a 45 degree angle in order to guide the nipple portion into the coupling. It was felt that these compliant features would also lead to enhanced manual operation of the coupling as well.

Coupling B had a very similar mechanism as coupling A. The narrow outer ring was pulled toward the flex hose until the breakout force of the coupling was overcome and the coupling was demated. Mating also occurred by aligning the coupling onto the nipple end and applying axial force until the coupling portion locked back into place.

Demating coupling C required depression of two detents, one on either side of a knurled aluminum ring. Once the detents were depressed, the aluminum ring would slide toward the flex hose and the coupling portion could be pulled away. Mating required aligning the coupling portion onto the nipple end and applying force axially until the detents engaged.

Coupling D had a lever actuated demating process. The coupling's lever was pushed toward the hard mounted, nipple end. When the lever was pushed beyond a certain point (approximately 45 degrees), demating automatically occurred. Mating required aligning the coupling and applying axial force onto the nipple end until the lever restored itself to the vertical position.

It is important to note that the task performed in this study does not represent the entire coupling process. The experimental task consisted of, in effect, the soft-latch phase of the coupling process where the coupling is mated or demated but the actual flow of fluid has not been affected. With each of these couplings, the flow of fluid would need to be

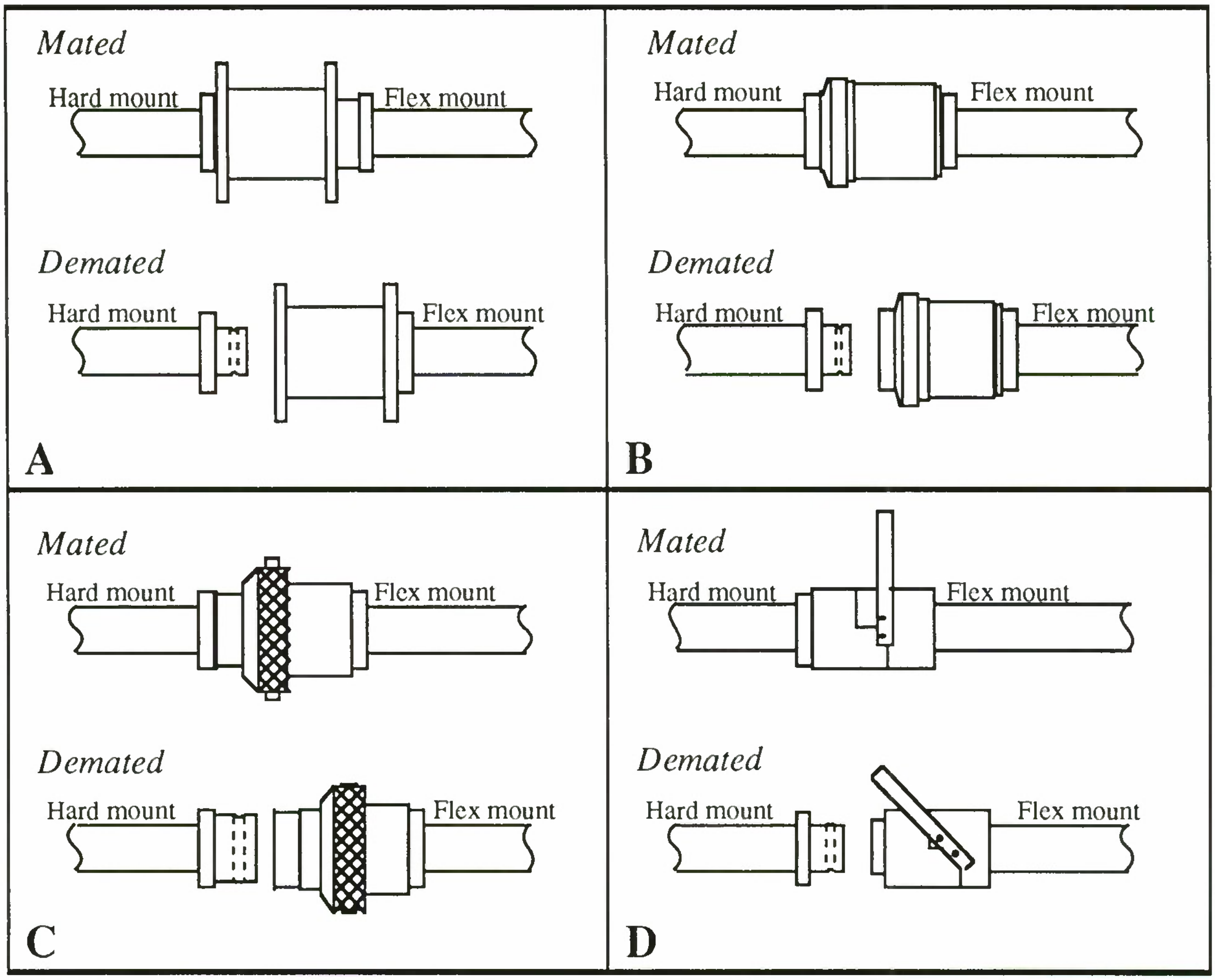


Figure 1. Schematic diagrams of the four couplings evaluated in this study.

turned on or off in an additional step not included in the task. That phase of the coupling process would involve the use of an added tool or modification to the end effector which would drive the coupling into the fully opened or closed position. Since that phase of the process has yet to be defined, it was of interest to the experimenters to evaluate the compatibility of the mating and demating components of the task which could be addressed at this time.

DESIGN

This study implemented a 2 modality (manual and telerobotic) by 4 coupling (couplings A, B, C, and D) within subjects design. The modality and the coupling sequence were counterbalanced as demonstrated in Table 1.

PROCEDURE

To begin each testing session, subjects were introduced to the purpose and procedure of the study as well as the basic layout of the cameras, task, and robotic system. Since subjects were already familiar with the operation of the robotic and viewing systems employed in the ROIL, no instruction was necessary regarding these aspects of the task. Subjects began the session by manipulating a coupling either manually or telerobotically depending on their particular counterbalancing sequence. Each coupling was demated and

mated three times in each modality. The experimenter kept performance time by means of a hand stopwatch and recorded those times on a data collection sheet where errors were logged as well. An error was counted only if the coupling portion was dropped, at which point the experimenter reset the coupling in the fully mated position. Following a set of three trials with each coupling, subjects filled out a short questionnaire with rating scales concerning workload, discomfort, as well as various task related issues. Once all the couplings had been completed, subjects filled out a final questionnaire for each modality where they rated the couplings in comparison to one another.

TABLE 1.

Counterbalancing sequence for couplings and modality across subjects (M = manual condition, T = telerobotic condition).

Subject	QD Coupling Sequence							
	1		2		3		4	
1	Coup. A		Coup. B		Coup. D		Coup. C	
	M	T	T	M	M	T	T	M
2	Coup. B		Coup. C		Coup. A		Coup. D	
	T	M	M	T	T	M	M	T
3	Coup. C		Coup. D		Coup. B		Coup. A	
	M	T	T	M	M	T	T	M
4	Coup. D		Coup. A		Coup. C		Coup. B	
	T	M	M	T	T	M	M	T

RESULTS AND DISCUSSION

Analysis of variance performed on the data showed there were clear trends in both the performance as well as the subjective data. Table 2 presents the group means for many of the performance and subjective measures. Due to the very few number of errors occurring in any of the trials, analysis of the error data resulted in no significant findings and is not presented in the table.

It was hypothesized that as a result of the compliant structures built into coupling A, demating and mating it would be faster than other couplings without these structures built into them. Data from the telerobotic trials showed that differences between performance time across the couplings was significant ($F(3,3) = 4.372$, $p < .05$). A Duncan's pairwise comparison performed on the data showed that the source of significance came largely from coupling C being significantly slower than all other couplings' performance time. Due primarily to the small variance in the manual condition, differences in performance time did not reach significance for these trials. A Duncan's pairwise comparison on these data, however, did show that performance time for coupling A was significantly faster than coupling C. As anticipated, it appears that for both modalities, coupling A was faster — in some cases significantly faster — to demate and mate than the other couplings.

TABLE 2.

Group means for performance and subjective measures.

Measure	Modality of Operation							
	Telerobotic				Manual			
	A	B	C	D	A	B	C	D
Perform. Time per Trial (sec.)	66	77	450	168	2.4	3.7	6.4	3.8
Overall Rating (1 to 7)	1.5	2.0	5.3	3.8	1.8	2.5	5.0	2.5
Grip Acceptability (1 to 7)	1.3	2.8	3.5	2.8	1.3	1.8	3.8	1.3
Mental Workload (1 to 10)	3.0	2.5	6.8	4.0	Not Addressed			
Phys. Discomfort (1 to 7)	2.5	1.5	3.8	2.5				

It was also felt that subjective reactions to the couplings would show preference for the custom coupling in both modalities. The overall rating data were collected on seven point scales with 1 corresponding to "completely acceptable" and 7 corresponding to "completely unacceptable." As shown in Table 2, these data revealed reliable differences, this time for both telerobotic as well as manual ratings. The data regarding the telerobotic preference revealed an $F(3,3) = 7.981$ with a $p < .01$. Pairwise comparisons showed that couplings A and B were rated significantly more acceptable than coupling C,

while coupling A was significantly more acceptable than coupling D. For the manual ratings the data showed that F (3,3) = 8.007, with $p < .01$. In this case pairwise comparisons indicated that coupling C was significantly less acceptable than all others. The comparable ratings attributed to couplings A and B appeared the result of their similar mechanisms and operation. The shape of the outer sleeve and coupling A's chamfering were all that varied between the two.

Using the same seven point scale described above, data regarding the acceptability of obtaining the proper grip did not reach significance for the telerobotic condition, although the pairwise comparisons did show that coupling A was rated significantly more acceptable than coupling C. For the manual condition this difference did reach significance, F (3,3) = 5.368, $p < .05$, with the comparisons among the means indicating that coupling C was significantly less acceptable than all three other couplings.

After the telerobotic trials, data were also collected on mental workload and physical discomfort. Data from a Modified Cooper-Harper mental workload rating scale reached significance, F (3,3) = 3.860, $p = .05$. The pairwise comparisons showed that couplings A and B were rated significantly less mentally taxing than coupling C. Data from either question addressing physical discomfort did not reach significance although the pairwise comparisons tended to show couplings A and B as less demanding than coupling C. These effects seemed the result of the rather straight-forward mechanism implemented on couplings A and B. Subjects only had to grab and pull to demate couplings A and B, while coupling C required depression of the detents on either side of the detention sleeve. This orientation was often very difficult to achieve with the robotic grippers, typically requiring repeated attempts before demate finally occurred. Issues of mental workload and physical discomfort were not addressed after the manual trials due to their very short duration.

CONCLUSION

Of the couplings included in this study, several design components were found to be of interest. With respect to the operation of the couplings, the various concepts resulted in differing reactions from the subjects.

Regarding the demate process, subjects felt coupling D included an attractive feature by requiring little force to demate, achieving it simply by forcing the lever over. However, maintaining control of the coupling portion after demate proved difficult for teleoperation, although somewhat easier for manual operation. Demating coupling C showed that depression of detents is a very delicate operation to perform with the telerobot and to some extent, to perform manually as well. Without some method of fixing the orientation of the detents, it is very difficult to engage both at the same time, particularly with the telerobot. This was compounded by the fact that the depression had to be combined with the axial force necessary to demate. Because of coupling B's small outer ring, demating was at times found to be clumsy with it as well. This was particularly the case for the telerobotic condition, but at times the manual condition was awkward as well.

Mating the couplings proved, on the whole, a far simpler process. Couplings B and D required close alignment which, when met, resulted in a very straightforward mating process. Coupling C incorporated a longer nipple portion to the coupling. This assisted operation in both modalities by helping to guide the coupling into the mated position when the axial force was applied.

While coupling A did appear the better design in this evaluation, there clearly were facets which could be improved. Although the large flanges on the outer sleeve assisted in mating, they also might allow the telerobot or EVA astronaut to accidentally bump or deactivate the coupling prior to full actuation. Also, the chamfering performed on the entry of the

coupling was perhaps angled too far. The 45 degree entrance guided the nipple portion into the coupling, but allowed sufficient misalignment such that the coupling often bound just prior to fully mating. Symetrics has recognized these concerns and has provided the ROIL a coupling addressing these issues by making two changes in the design. New shorter flanges still allow necessary support for the axial forces required, but greatly reduce the likelihood of accidental deactivation. The entry to the coupling was also chamfered to approximately 30 degrees rather than 45. This assisted in guiding the nipple into the coupling but reduced the potential for binding by lessening the amount of misalignment possible.

The purpose of this study was not to conceive the final coupling design. Rather, it was intended as a step along an iterative process. The newly modified coupling will be included in a series of further controlled, as well as subjective, evaluations. This is part of ongoing work in the ROIL designed to enhance the overall interface by improving design at both the teleoperator and telerobot ends of the system.

ACKNOWLEDGEMENTS

Support for this investigation was provided by the National Aeronautics and Space Administration through Contract NAS9-17900 to Lockheed Engineering and Sciences Company.

The authors would like to acknowledge John Calvin and Steve Calvin of Symetrics Incorporated as well as Ray Anderson of MacDonnell Douglas Space Systems Company for their assistance in providing materials and information crucial to this evaluation.

REFERENCES

1. Hannaford, B., Wood, L., Guggisberg, B., McAffee, D., and Zak, H. *Performance evaluation of a six-axis generalized force-reflecting operator.* (JPL Publication 89-18). Pasadena, California: NASA Jet Propulsion Laboratory.

2. National Aeronautics and Space Administration (1989). *Interface design considerations for robotic satellite servicers.* (NASA Report JSC-23920). Satellite Servicing Systems Project Office.

3. Newport, C. (1989, June). *Application of subsea robotics maintenance experience to satellite servicing.* Paper presented at the Satellite Services Workshop IV, NASA Johnson Space Center, Houston, Texas.

THE EFFECTS OF SPATIALLY DISPLACED VISUAL FEEDBACK ON REMOTE MANIPULATOR PERFORMANCE

Randy L. Smith and Mark A. Stuart
Lockheed Engineering and Sciences Company

Research directed by Jay Legendre, Manager, Remote Operator Interaction Lab, NASA JSC.

INTRODUCTION

Telerobotics will be heavily used for the assembly, maintenance, and servicing of NASA's Space Station *Freedom*. The visual system may well be the single most important source of information for the operator of the various telerobot systems that will be used. When performing a remote manipulation task, the operator can view the remote scene either by looking through a window, or with the use of cameras. For most of the tasks that will be performed on the Space Station, a direct view of the work area will either not be available, or will not provide the necessary visual cues for teleoperation. Therefore, cameras will provide the primary mode of feedback to the operator concerning manipulator position, orientation, and rate of movement.

Operators normally use the body of the manipulator as a reference point when making control inputs, but, if the Space Station's external cameras are placed such that the camera view is not normal with respect to the manipulator (normal refers to placement approximately behind the shoulder of the manipulator arm) then the visual feedback will be spatially displaced. At a fundamental level, displacement refers to there not being a one-to-one spatial correspondence between control inputs and perceived motion (either directly perceived through a window view or perceived on monitors). Spatial displacement is an unfortunate consequence of attempts to provide visual information to the operator when the camera placement is not normal and it should be avoided if at all possible. If displaced visual feedback is presented to the operator, system performance can be seriously degraded due to operator disorientation. This is important for Space Station *Freedom* telerobotic tasks because it is possible that cameras may be placed on Station structure such that the human operator receives displaced visual feedback.

If control inputs are referenced to the body of the manipulator (analogous to the "world" mode in industrial robotics) the following descriptions can be made. Spatially displaced feedback can take on four different forms: *angular displacement,* where the reference point is displaced horizontally or vertically (see Figure 1 for a depiction of horizontally displaced angular feedback and Figure 2 for vertically displaced angular feedback); *reversal displacement*, where the camera is facing the arm instead of being placed behind it (see Figure 3 for a depiction of reversal); *inversion-reversal displacement*, where the camera is upside down and is facing the arm (see Figure 3); and *inversion displacement*, where the camera is upside down with respect to the manipulator arm (see Figure 3).

It has been suggested that these spatial displacements adversely affect operator performance to varying degrees. The literature states that <u>direct</u> manipulation tasks take on progressively more disturbance, with angular displacement being the least disruptive and inversion displacement being the most disruptive (Smith and Smith, 1962). Direct manipulation is where a person manipulates an object with their bare hands or with a simple, rigid tool such as a stylus or screwdriver. A remote manipulation task is where the person manipulates a mechanism (e.g., hand controller) which transfers the operator's motions to a remotely placed mechanical device. The actual manipulation of the object is spatially removed or distant from the operator.

Early studies on spatial displacement were conducted by Helmholtz, Kohler, Smith and

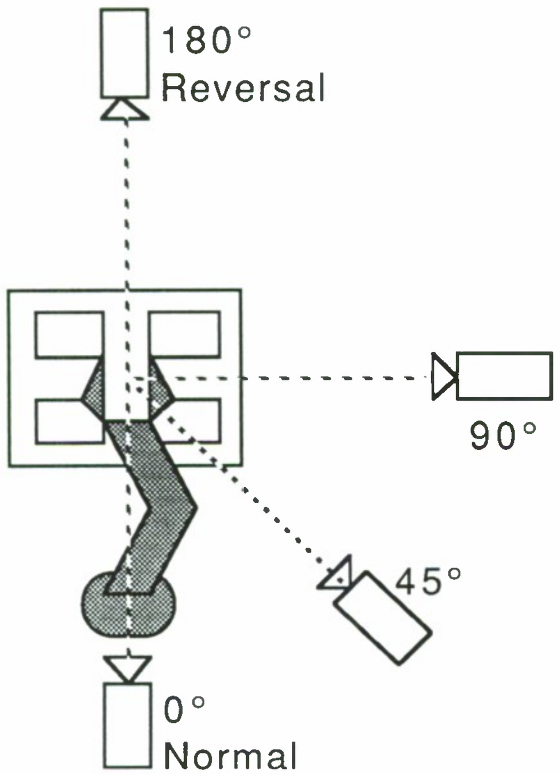


Figure 1. Angularly displaced feedback, horizontally displaced about the manipulator (top view shown).

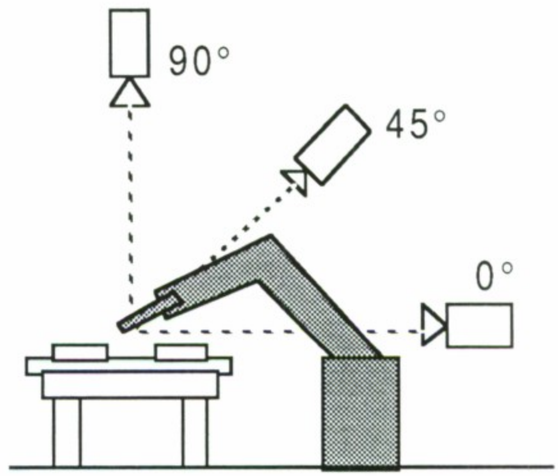


Figure 2. Angularly displaced feedback, vertically displaced about the manipulator (side view shown).

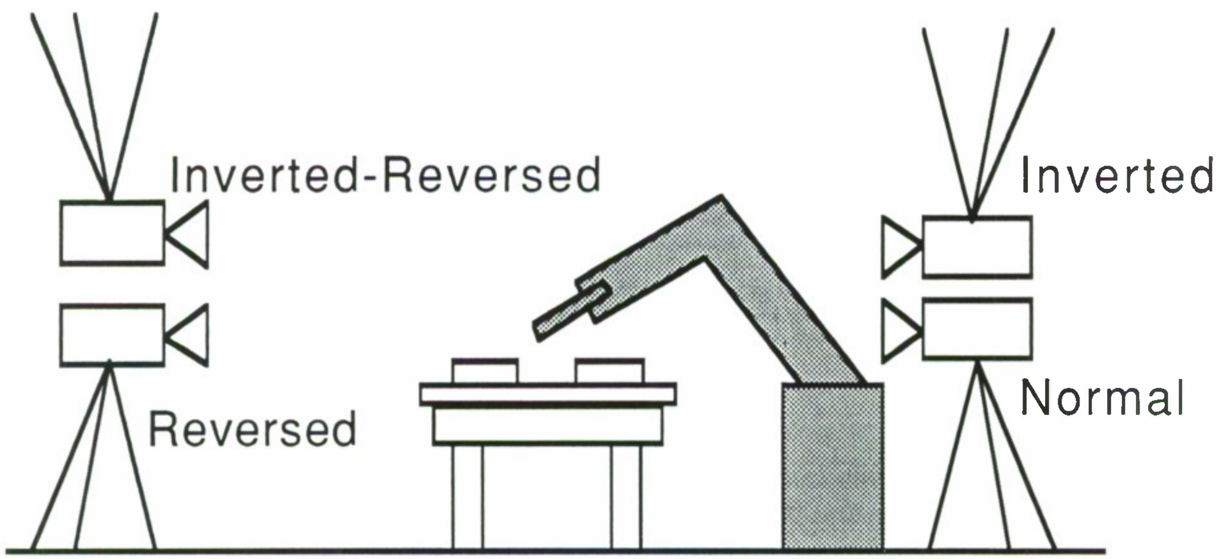


Figure 3. Spatially displaced feedback -- normal, reversal, inversion, and inversion-reversal (side-view shown).

Smith, et al. Smith and Smith (1962) were interested in perceptual-motor integration, specifically the effects of spatial and temporal displacements of the visual feedback of motion. According to Smith and Smith (1987) their work was incorporated into the neurogeometric organization of behavior in which "space perception and visually controlled movement are learned, the nature and degree of learning are determined by the nature and degree of spatial compliance between muscular control and sensory input." There have been numerous studies of viewing systems for telerobotic systems; however, it is difficult to draw coherent and generalized conclusions from them (Crooks, Freedman, and Coan, 1975; Horst, Rau, LeCocq, and Silverman, 1983; Chu and Crooks, 1980; Clarke, Hamel, and Draper, 1983; Bodey and Cepolina, 1973; Huggins, Malone, and Shields, 1973; Onega and Clingman, 1973; Fornoff and Thornton, 1973; and Clarke, Handel, and Garin, 1982).

OBJECTIVES

One objective of this investigation was to quantify whether the above mentioned results from the literature hold true for remote manipulation tasks performed with a remote manipulator arm. It was also of interest to informally evaluate how a direct view of the worksite compares to a normal camera-aided view of the worksite. This secondary evaluation was an attempt to determine if the results obtained in a previous evaluation (Smith, 1986) of remote manipulator operators, who had both direct and normal camera views, could be replicated. The present investigation examined operators performing a remote manipulation task while exposed to the following different viewing conditions:

- direct view of the work site (baseline condition)
- normal camera view (zero-degree displacement)
- reversed camera view (180-degree displacement)
- inverted/reversed camera view
- inverted camera view

METHOD

Data were collected from subjects as they performed a remote manipulation task while exposed to the different viewing conditions. All six subjects used the five viewing conditions.

SUBJECTS

Six volunteer subjects were used in this evaluation. All were experienced in the operation of the Kraft robotic system.

APPARATUS

Testing took place in the Man-Systems Telerobotics Laboratory at NASA's Johnson Space Center (JSC).

A Kraft Telerobotics force-reflecting 6 degree-of-freedom master-slave remote manipulator was used to perform the remote manipulation task. A Javelin CCD color camera and a Mitsubishi 20-inch color video monitor were used to present the camera views to the subjects. The camera in each position was exactly 10 feet 2 inches from the remote manipulation task, with the focus and zoom controlled. The direct view was from a distance that was controlled so that the visual angle subtended at the eye by the task piece was approximately the same for both camera and direct views.

The task consisted of grasping and moving six pyramid-shaped wooden blocks so that they could be dropped inside a metal box located six inches from the blocks on the taskboard. This particular task was selected because it is functionally similar to multi-axis translation and alignment tasks which will be performed by the telerobots on Space Station *Freedom.*

Each subject sat behind a barrier for the camera-aided viewing conditions so that a direct view of the worksite was not possible. For the direct viewing condition, the subjects faced the Kraft manipulator arm with a zero-degree displacement view.

VARIABLES

The independent variable in this study was the different camera viewing conditions (normal, image reversal, image inversion/reversal, and image inversion). Note that the noncamera viewing condition (direct view) was not an independent variable, but was only used as a baseline measure. This evaluation used a one-factor repeated measures design — all subjects were exposed to all levels of the independent variable used. The dependent variable in this evaluation was task performance time.

PROCEDURES

Subjects were instructed to perform the remote manipulation task quickly and accurately. All subjects performed the manipulation task with the direct view first. This served as the baseline condition by which the performance times for all the other viewing conditions could be compared. Each subject then performed the manipulation task for each of the four different camera-viewing conditions. Each subject performed the task a total of five times. The order in which subjects were exposed to the four different camera views was counterbalanced to control for order effects. Task performance times were collected throughout the test sessions.

RESULTS AND DISCUSSION

The task completion time data were collected and summarized. The average performance times (in minutes) are summarized as follows:

Direct (baseline)	0.59
Normal	1.20
Inverse/Reverse	5.00
Reverse	6.02
Inverse	9.51

The task completion times were then statistically analyzed with a repeated measures analysis of variance. It was determined from the ANOVA that the main effect of the viewing conditions, $F(3,15) = 7.72$, $p < 0.05$, was statistically significant. Because of this result, a Newman-Keuls pairwise comparison test was then administered to the data. It was revealed that the performance times for the inverted camera view were significantly ($p < 0.05$) worse than all of the other viewing conditions. This analysis also revealed that the reversed viewing condition was significantly worse than the normal viewing condition. The performance times for the inverted camera viewing condition were also significantly worse than the normal viewing condition performance times at $p < 0.01$.

It was also of interest to compare the performance of subjects under the direct-viewing condition to the performance of subjects under the normal camera-viewing condition. It was hoped that a statistically valid comparison could be made between these two viewing conditions, but since all subjects performed the direct-viewing condition first and the normal-viewing condition either second, third, fourth, or fifth out of all viewing conditions used in this study, then the results for this analysis could well be contaminated by the effect of differential amounts of training. A valid statistical comparison could not be made between these two viewing conditions because the experimental design in this study was used to counterbalance four different viewing conditions, not two. An informal comparison was made for the sake of general interest. This informal comparison involved partitioning the normal viewing-condition data from the rest of the camera-viewing data. These data and the direct-viewing data were analyzed with a *t*-test. This data analysis revealed that the task performance times for the normal viewing condition were significantly slower ($p < 0.05$) than for the direct viewing condition. It is recommended that further studies conduct an analysis of these two viewing conditions under proper experimental conditions so that an accurate assessment can be attained.

It is clear from the results of this study that spatially displaced visual feedback adversely affects remote manipulation performance. To get an indication of how views of the remote manipulator through a CRT monitor change with respect to hand controller movements for the four types of visual feedback studied in this

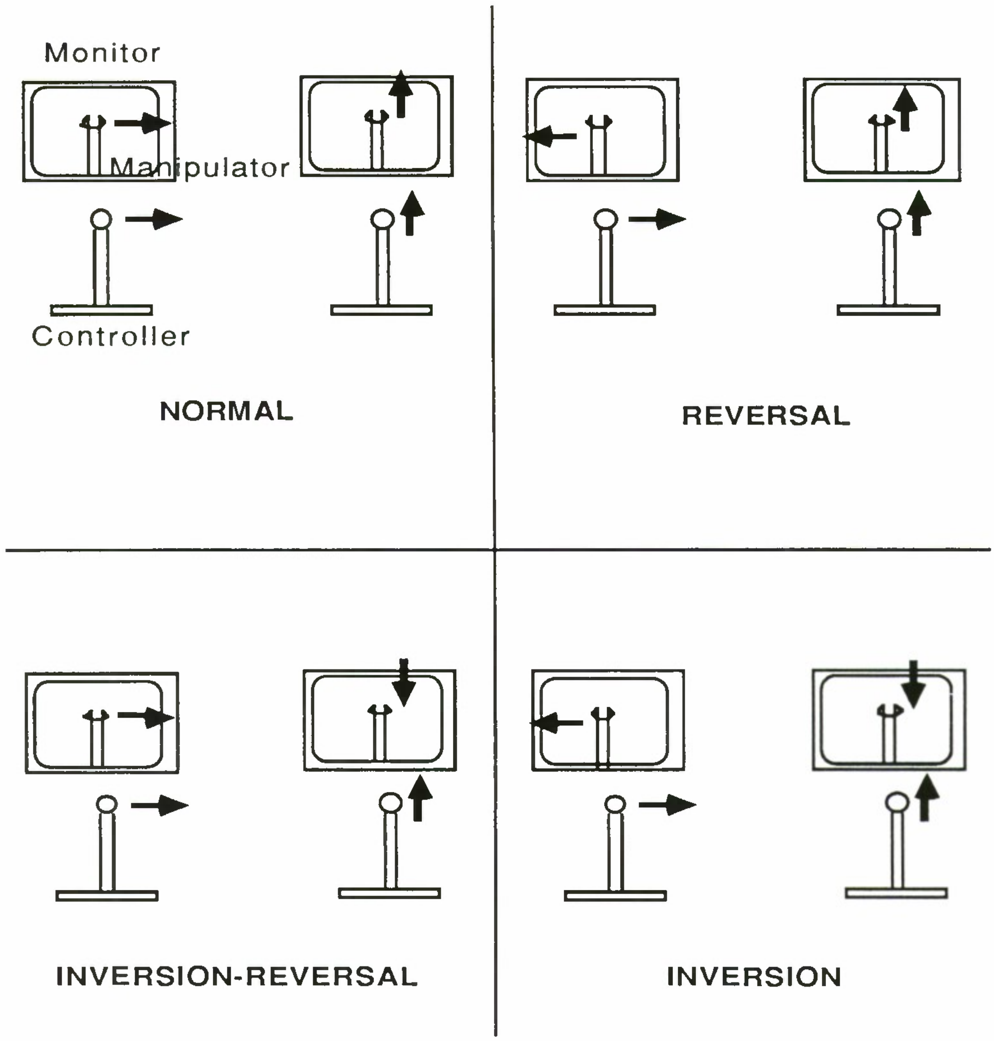


Figure 4. Viewed manipulator movements with respect to controller movements for four types of visual feedback.

evaluation, refer to Figure 4. In this figure, the object in the monitor is the remote manipulator. The arrows indicate the direction that the manipulator will move in the video image with respect to the specific hand controller movement for the four different camera placements. For example, in the inverted visual feedback condition, a rightward movement of the hand controller will result in the monitor image of the manipulator moving leftward and an upward hand controller movement will result in the monitor image of the manipulator moving downward.

The results obtained were not quite as would be expected based upon the previously mentioned studies of camera-aided viewing of direct manipulation tasks. The difference observed in this evaluation was that, in ranking the four viewing conditions, the reversed camera view was ranked third while the literature stated that the inversion/reversal was third. The reversed-viewing condition not only took over a minute longer, on the average, to complete than the inversed/reversed condition, but it was also significantly worse than the normal viewing condition performance

time. The differences obtained in this evaluation could be due to the fact that the remote manipulation task used in the present study involved the use of axes of movement different from those involved in the direct manipulation tasks reported in the literature. The axis of movement, the quantity of movement per axis, and the type of displacement are interrelated in a fashion that probably affects performance times; however, quantitative determination of these interrelationships is beyond the scope of this preliminary evaluation. The differences obtained could simply be due to the fact that the sited direct manipulation results were based upon data gathered from many different studies while the remote manipulation data came from only one study. More remote manipulation studies will need to be conducted before this conclusion can be made.

This study did informally replicate the results of the previously mentioned study by Smith (1986) which found, among other things, that performance with a normal camera view of the worksite is significantly worse than performance with a direct view. This result is no doubt partially due to the lack of binocular disparity that accompanied the camera viewing conditions.

CONCLUSIONS

The results of this evaluation have important implications for the arrangement of remote manipulation worksites and the design of workstations for telerobot operations. This study clearly illustrates the deleterious effects that can accompany the performance of remote manipulator tasks when viewing conditions are less than optimal. Future evaluations should emphasize telerobot camera locations and the use of image/graphical enhancement techniques in an attempt to lessen the adverse effects of displaced visual feedback. For a further discussion of the effects of perturbed sensory feedback see Smith, Smith, Stuart, Smith and Smith (1989).

An important finding in this evaluation is the extent to which results from previously performed direct manipulation studies can be generalized to remote manipulation studies. Even though the results obtained were very similar to those of the direct manipulation evaluations, there were differences as well. This evaluation has demonstrated that generalizations to remote manipulation applications based upon the results of direct manipulation studies are quite useful, but they should be made cautiously.

ACKNOWLEDGEMENTS

Support for this investigation was provided by the National Aeronautics and Space Administration through Contract NAS9-17900 to Lockheed Engineering and Sciences Company.

REFERENCES

1. Bodey, C. E., and Cepolina, F. J. (1973). Shuttle-attached manipulator system requirements. In E. Heer (Ed.), *Remotely Manned Systems* (pp. 77-83). Pasadena, CA: California Institute of Technology.

2. Chu, Y., and Crooks, W. (1980). Man-machine communication in computer-aided teleoperator control systems. In *Proceedings of the Human Factors Society 24th Annual Meeting* (pp. 54-58). Santa Monica, CA: Human Factors Society.

3. Clarke, M., Hamel, W., and Draper, J. (1983). Human factors in remote control engineering development activities. In *Proceedings, 31st Conference on Remote Systems Technology* (pp. 8-16). Gatlinburg, TN: American Nuclear Society.

4. Clarke, M., Handel, S., and Garin, J. (1982). A comparison of color and black and white television for remote viewing in nuclear facilities. In *Transactions of the American Nuclear Society* (p. 239). Washington, DC: American Nuclear Society.

5. Crooks, W., Freedman, L., and Coan, P. (1975). Television systems for remote manipulation. In *Proceedings of the Human Factors Society 19th Annual Meeting* (pp. 428-435). Santa Monica, CA: Human Factors Society.

6. Fornoff, H., and Thornton, W. G. (1973). Experimental evaluation of remote manipulation systems. In E. Heer (Ed.), *Remotely Manned Systems* (pp. 43-61). Pasadena, CA: California Institute of Technology.

7. Horst, R., Rau, P., LeCocq, A., and Silverman, E. (1983). Studies of three-dimensional viewing in teleoperated systems. In *Proceedings, 31st Conference on Remote Systems Technology* (pp. 30-34). Gatlinburg, TN: American Nuclear Society.

8. Huggins, C. T., Malone, T. B., and Shields, N. L. (1973). Evaluation of human operator visual performance capability for teleoperator missions. In E. Heer (Ed.), *Remotely Manned Systems* (pp. 337-350). Pasadena, CA: California Institute of Technology.

9. Onega, G. T., and Clingman, J. H. (1973). Free-flying teleoperator requirements and conceptual design. In E. Heer (Ed.), *Remotely Manned Systems* (pp. 19-41). Pasadena, CA: California Institute of Technology.

10. Smith, K. U., and Smith, W. M. (1962). *Perception and Motion.* W. B. Saunders and Company: Philadelphia.

11. Smith, R. L. (1986). *The effects of televised angular displacement of visual feedback on a remote manipulation task.* Unpublished Masters Thesis, Texas A&M University.

12. Smith, R. L. (1988). Human Visual Requirements for Control and Monitoring of a Space Telerobot. In W. Karwowski, H. R. Parsaei, and M. R. Wilhelm (Eds.), *Ergonomics of Hybrid Automated Systems I*, Elsevier: Amsterdam, pp. 233-240.

13. Smith, T. J., and Smith, K. U. (1987). Feedback-Control Mechanisms of Human Behavior. In G. Salvendy (Ed.), *Handbook of Human Factors.* John Wiley & Sons: New York.

14. Smith, T. J., Smith, R. L., Stuart, M. A., Smith, S. T., and Smith, K. U. (1989). Interactive Performance in Space - The Role of Perturbed Sensory Feedback. *3rd International Conference on Human-Computer Interaction.* Boston: Massachusetts, September 18-22.

SIMULATION OF THE HUMAN-TELEROBOT INTERFACE ON THE SPACE STATION

Mark A. Stuart and Randy L. Smith
Lockheed Engineering and Sciences Company

Research directed by Jay Legendre, Manager, Remote Operator Interaction Lab, NASA JSC.

INTRODUCTION

The Space Station is a NASA project which, when completed in the mid-1990s, will function as a permanently manned orbiting space laboratory. A part of the Space Station will be a remotely controlled Flight Telerobotic Servicer (FTS). The FTS, a project led by NASA's Goddard Space Flight Center, will be used to help assemble, service, and maintain the Space Station and various satellites. The use of the FTS will help ensure the safety and productivity of space-based tasks normally accomplished by astronauts performing outside the pressurized spacecraft. For the short-term, control of the FTS will be dependent primarily on the human operator. Since the human operator will be a part of the telerobotic system, then it is important that the human-telerobot interface be well-designed from a Human Factors perspective. It is critical that the components of this interface be designed so that the human operator's capabilities and limitations are best accommodated for within the structure of specific task requirements. To emphasize the importance of a well-designed human-telerobot interface, one study found that simply the selection of an appropriate control device, based upon the operator's capabilities and the requirements of the task, can more than double the productivity of the telerobotic system (O'Hara, 1986).

With the system development process becoming more complex and expensive, more emphasis is being placed on the evaluation of systems during early stages of the development cycle. The design of systems that include human operators is especially complex because determining overall systems performance is dependent upon the interaction of the human operator, hardware components, and software components (Chubb et al., 1987). Adequately evaluating the performance of a system during the design cycle is becoming increasingly more difficult when using the static evaluation tools traditionally available to the Human Factors engineer, such as job and task analysis (Geer, 1981). It is becoming more common for systems developers to use computer simulation as a design tool instead of hardware models (Gawron and Polito, 1985) and for Human Factors engineers to use computer simulation to enhance the use of static evaluation tools. This is because more sophisticated analysis tools are needed that will allow a controlled evaluation of the human operator/hardware components/software components interaction (Chubb, et al., 1987).

This paper will cover the various uses of simulation, the elements of the human-telerobot interface, and how simulating the human-telerobot interface on the Space Station will result in a better designed system. Before focusing the discussion specifically to the simulation of the human-telerobot interface, it will be useful to briefly define simulation and to cover the major uses of system simulation—independent of the type of system that is being simulated. There will then be a discussion of the areas of the human-telerobot interface and how simulation can contribute to a better designed user interface from a Human Factors perspective.

USES OF SIMULATION

Simulation is the process of imitating or duplicating the actions or processes of some system in a controlled environment (Arya, 1985). Emphasis should be placed on the word "controlled." System simulation, either hardware, computer, or a combination of the two, has been used for decades. This paper

will describe four major uses of simulation. One use of simulation is to study the effectiveness of various hardware/software components on overall system's performance. The advantages of using simulation within this context are *cost* — it is cheaper to simulate a system than it is to build one; *time* — simulating a system is usually faster than building it; *feasibility* — because of the size and complexity of some systems, it is not possible to evaluate them in the real world, therefore, simulation serves the function of systems verification; *safety* — some systems operate in dangerous environments and can only be evaluated safely with the use of simulation; and *prediction* — with the use of simulation, a system's performance and processes can be speeded up so that future behavior can be predicted (Arya, 1985).

A second use of simulation is to study the effects of various hardware/software components on simulated human performance. This approach utilizes mathematical models of human performance to assist the simulation process. In this, as well as, the approach mentioned above, man-in-the-loop is not a part of the evaluation.

A third use of simulation is to study the effects of various hardware/software components on actual human performance. This approach can be taken in an attempt to match systems components and operator capabilities and limitations in order to ensure optimal systems and operator performance. This approach can be taken in an attempt to add greater fidelity, and thus, external validity to the data that are gathered in the analysis.

The last use of simulation to be addressed in this paper is to train operators to eventually use a real-world system. The major benefits of simulation as a training aid are in the areas of *scheduling* — training is not affected by weather or the need to perform operational missions; *cost* — simulator training is significantly less expensive than prime system training; *safety* — reduces the exposure of operators and the prime system to the hazards of the operating environment; *control of training conditions* — control of environmental and human interaction conditions that may be a part of the operating environment; *learning enhancement* — system malfunctions and environmental conditions can be included in the training; and *performance enhancement* — inclusion of critical missions that are difficult to train for in the real world (Flexman and Stark, 1987).

As the above list indicates, simulation has significant usage as an aid in the development of pre-existent systems. It can have even greater significance in the design and development of novel pre-existent systems — systems that have never existed before and where few direct comparisons to existent systems can be made. The human-telerobot system that will be used on the Space Station is such a novel system.

Even though industrial robots and teleoperators are heavily used in such areas as the nuclear industry and in underwater activities, there are major differences between these applications and the telerobot system to be used on the Space Station — one of these being the zero-gravity factor. There is also a limited number of direct comparisons which can be made from the Remote Manipulator System (RMS) used on the Space Shuttle and from the proposed telerobot system. The review of the literature concerning these systems has provided answers to some important design issues, but there are major limits to how far these data can be generalized to the human-telerobot interface on the Space Station. Laboratory evaluation of the effects of various hardware and software components on operator performance can, of course, provide answers and guidance, but, perhaps greater fidelity can be attained with the use of simulation.

It is thus proposed that the use of simulation in the design and development of the human-telerobot interface on the Space Station will be very beneficial. Simulation should serve as an aid in the selection and design of hardware and software components to ensure maximum, error-free performance. Simulation should be worthwhile especially for its ability to simulate the effects of zero gravity on performance. Operator performance at manipulation tasks

while in a one-gravity environment may well not be generalizable to weightless states. Simulation of the interface should also have the benefit of helping engineers to detect flaws in the design of components of the interface which would adversely affect system and/or operator performance. It is obviously important that any mistakes of this type be detected early and far before the design is finalized or manufacture of the system has occurred.

INFORMATION NEEDS OF THE OPERATOR

There are three broad areas of the human-telerobot interface where simulation can be of assistance: operator information needs, control devices, and workstation layout. These three areas are listed in Table 1. The information needs of the operator will vary depending upon the tasks to be performed. The operator will need information concerning the location and orientation of the telerobot in space, the health status of the telerobot, visual feedback from the viewing system, the status of any transportation devices, the status of the workpiece, and the status of the hardware in the control workstation.

Regarding visual feedback, the visual system may well be the single most important source of information for the operator (Smith and Stuart, 1988). Some of the issues related to the visual system are concerned with camera position and number, the spatial orientation of the image presented to the operator, and monitor type, placement, and number. For example, when performing a remote manipulation task in real time, the operator can view the remote scene either by looking through a window, or with the use of cameras. For most of the tasks that will be performed in space, a direct view of the working area will either not be available, or will not provide the necessary visual cues for teleoperation. Therefore, cameras will provide the primary mode of feedback to the operator concerning manipulator position, orientation, and rate of movement. Operators normally use the body of the manipulator as a reference point when making control inputs, but if the Space Station's external cameras are placed such that the camera view is not normal to the manipulator (normal refers to placement behind the shoulder of the arm), then the visual feedback will be spatially displaced. Spatial displacement is an unfortunate consequence of attempts to provide visual information to the operator when the camera placement is not normal and it should be avoided if at all possible.

TABLE 1.

Three areas of the human-telerobot interface

1. Information needs of the operator
   - Location of telerobot
   - Status of transportation devices
   - Status of workpiece
   - Status of workstation
   - Force feedback
   - Visual feedback
     - Camera position and number
     - Spatial orientation of image
     - Monitor type, placement, number
     - Illumination

2. Control devices considered
   - Miniature master controllers
   - 3 or 6 degree-of-freedom hand controllers
   - Exoskeleton controllers
   - Head-slaved controllers
   - Dedicated switches
   - Programmable display pushbuttons
   - Voice command systems
   - Computers

3. Telerobot workstation
   - Hardware layout
   - Software layout

Spatially displaced feedback can take on different forms: *angular displacement,* the reference point is displaced horizontally within the sagittal plane or vertically within the median plane; *reversal* is facing the arm instead of being placed behind it; *inversion-reversal* is upside down and is facing the arm; and *inversion*, the camera is upside down with

respect to the manipulator arm. The image can also be displaced *temporally* — there are time delays in which the operator receives the visual feedback, as well as *size distorted* — the image is enlarged or reduced from its actual size. These spatial displacements adversely affect operator performance to varying degrees. Generally, they take on progressively more disturbance with angular displacement being the least disruptive and inversion displacement being the most disruptive. Temporal displacement interrupts the intrinsic temporal patterning of motion and causes severe disruptions in behavior. Much effort should be extended to prevent its occurrence. Size distortions generally do not affect performance to a great extent (Smith and Smith, 1962).

Other visual system issues include how an operator will use multiple views of the task area and how operators can best use non-stereoscopic cues to depth perception. Computer simulation of various task scenarios with human operators working within various hardware and software mockups, including sophisticated scene generation techniques, can serve as an aid in determining what types of information are needed and what types of information presentation enhancements should be used at various points within the sequence of task performance. An example of an information enhancement technique that simulation can investigate is the use of real-time moving graphics displays designed to help operators maintain their orientation while performing under potentially visually disorienting conditions. Other screen-viewing techniques should be investigated with the use of simulation in an attempt to avoid operator disorientation while performing manipulation tasks.

CONTROL DEVICES

Control devices will be used to control such things as telerobot activation, position, manipulators, end effectors, rate of movement, and the viewing system. Control devices being considered include manipulator controllers such as miniature master controllers with direct position control, 3 or 6 degree-of-freedom hand controllers using rate or force inputs, exoskeleton controllers using various position sensors to detect human arm configurations, head-slaved control, dedicated switches, programmable display pushbuttons, voice-commanded systems, and computer displays with cursor-control devices which allow menu selections. Control device selection is important because it affects operator performance, workload, and preference. Computer simulated scenarios could be linked to actual controllers' use to determine their effects on operator performance across different manipulation tasks.

WORKSTATION DESIGN

The telerobot workstation consists of hardware elements, their interfaces, and the software that will allow the hardware to be used. The workstation is the point where the information and control inputs are made available to the operator. Just as with the selection of control devices, the workstation should be logically and functionally laid out to optimize operator performance and preference while minimizing workload and error rates. Again, simulation can help to determine optimal workstation layouts. A simple means of simulating the workstation layout is through the use of computer prototyping, but it is recommended that large-scale simulation be used as a means of designing and evaluating the telerobot workstation.

CONCLUSIONS

Many issues remain unresolved concerning the components of the human-telerobot interface mentioned above. It is then critical that these components be optimally designed and arranged to ensure, not only that the overall system's goals are met, but that the intended end-user has been optimally accommodated. With sufficient testing and evaluation throughout the development cycle, the selection of the components to use in the final telerobotic system can promote efficient, error-free performance. It is recommended that whole-system simulation with full-scale mockups be used to help design the human-telerobot interface. It is contended that the use of simulation can facilitate this design and evaluation process. The use of simulation can also ensure

that the hardware/software components have been selected to best accommodate the astronaut, instead of the astronaut having to make performance accommodations for the hardware/software components that have been selected.

As was mentioned above, there are other advantages to simulating the human-teleoperator interface than simply serving as an aid in the selection and design of hardware/software components so that operator performance is optimized. Systems developers can also use the simulation system to test whether or not hardware components meet overall systems goals, and the simulation system can be used for subsequent training of the astronauts who will use the actual system.

ACKNOWLEDGEMENTS

Support for this investigation was provided by the National Aeronautics and Space Administration through Contract NAS9-17900 to Lockheed Engineering and Space Company.

REFERENCES

1. Arya, S. (1985). Defining various classes of simulators and simulation. In J. S. Gardenier (Ed.), *Simulators* (pp. 36-38). La Jolla, California: Simulation Councils, Inc.

2. Chubb, G. P. Laughery, Jr., K. R., and Pritsher, A. A. B. (1987). Simulating manned systems. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 1298-1327). New York, New York: John Wiley and Sons.

3. Flexman, R. E., and Stark, E. A. (1987). Training simulators. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 1012-1038). New York, New York: John Wiley and Sons.

4. Gawron, V. J., and Polito, J. (1985). Human performance simulation: combining the data. In J. S. Gardenier (Ed.), *Simulators* (pp. 61-65). La Jolla, California: Simulation Councils, Inc.

5. Geer, C. W. (1981). *Human engineering procedures guide* (AFAMRL-TR-81-35). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.

6. O'Hara, J. M. (1986). *Telerobotic work system: Space-station truss-structure assembly using a two-arm dextrous manipulator* (Grumman Space Systems Report No. SA-TWS-86-R007). Bethpage, NY: Grumman Space Systems.

7. Smith, K. U., and Smith, W. M. (1962). *Perception and motion: An analysis of space structured behavior.* Philadelphia, Pennsylvania: W. B. Saunders Company.

8. Smith, R. L., and Stuart, M. A. (1988). Telerobotic vision systems: The human factor. In *Proceedings of the Instrument Society of America (ISA) Conference* (in press).

Ergonomics

An astronaut's applied force is measured to determine human capability to perform tasks in space. The CYBEX dynanometer, which is used to evaluate the skeletal muscle strength, power, and endurance of astronauts and astronaut candidates, is used in the Weightless Environment Training Facility (WETF) to simulate zero-gravity.

QUANTITATIVE ASSESSMENT OF HUMAN MOTION USING VIDEO MOTION ANALYSIS

John D. Probe
Lockheed Engineering and Sciences Company

Research directed by Michael Greenisen, Former Manager, Anthropometry and Biomechanics Lab, NASA JSC.

INTRODUCTION

In the study of the dynamics and kinematics of the human body a wide variety of technologies has been developed. Photogrammetric techniques are well documented and are known to provide reliable positional data from recorded images. Often these techniques are used in conjunction with cinematography and videography for analysis of planar motion, and to a lesser degree three-dimensional motion. Cinematography has been the most widely used medium for movement analysis. Excessive operating costs and the lag time required for film development, coupled with recent advances in video technology, have allowed video based motion analysis systems to emerge as a cost effective method of collecting and analyzing human movement. The Anthropometric and Biomechanics Lab at Johnson Space Center utilizes the video based Ariel Performance Analysis System to develop data on shirtsleeved and space-suited human performance in order to plan efficient on-orbit intravehicular and extravehicular activities.

VIDEO BASED MOTION ANALYSIS

The Ariel Performance Analysis System (APAS) is a fully integrated system of hardware and software for biomechanics and the analysis of human performance and generalized motion measurement. Major components of the complete system include the video system, the AT compatible computer, and the proprietary software.

VIDEO SUBSYSTEM

The video system consists of commercial quality 112-inch VHS format portable video cameras. They are used to record motion sequences for subsequent analysis. A minimum of two cameras are required for full three-dimensional analysis. A high quality VCR and monitor are used for the display and digitizing of videotaped sequences. The playback unit accommodates standard VHS videotapes recorded from any standard video source to allow high precision freeze-frame video imaging with accurate single frame advance and reverse as well as a variable speed search capability.

HARDWARE SUBSYSTEM

An AT compatible computer is the primary component of the analysis system. The computer uses a combination of a frame grabber and a VCR controller board to digitize from the playback unit. The APAS captures video images from video tape and imports them into the computer memory. The operator can then digitize the desired sequence by positioning a cross-hair cursor over the joint center of interest and recording the coordinates of this point by pressing a button on the mouse. Digitization of the joints on the first frame is performed completely by the operator. For subsequent frames, the point locations from previous frames are used to predict the positions of each point on the current frame. This significantly reduces the time required to digitize a sequence. Additionally, since the computer stores a digital image, the analysis frame can be contrasted, enhanced, or filtered for clarification.

SOFTWARE SUBSYSTEM

An extensive integrated software package makes up the third component of the APAS. For ease of operation, the software has been highly structured and modularized. Each module is designed to perform a particular function and is completely menu driven. A

brief functional description of each module is listed below.

Performance analysis always begins with the *digitizing* module. This module allows video images to be converted to body joint location coordinates in the computer. These digitized locations are saved for subsequent conversion to true image space location.

The *transformation* module converts digitized video data into true two or three-dimensional image data using an algorithm called direct linear transformation (1). If a single camera view is used, the resulting image is two-dimensional. If two or more camera views have been used, the resulting image is three-dimensional.

The *smoothing* module removes small random digitizing errors from the computed image coordinates. At the same time, it computes body joint velocities and accelerations from the smoothed joint coordinates. The operator may choose any of three different smoothing functions: cubic spline, digital filter, and polynomial smoothing. The APAS utilizes, as the smoothing method of choice, a modified cubic spline smoothing algorithm formulated by Reinsch (2). In addition, the user may control the amount of smoothing applied to each joint to insure that smoothing does not distort the digitized data.

The *viewing* module is used to examine image data in "stick figure" format. Viewing options include single frame, multiple frame, and animated images. Three-dimensional images may be rotated to allow viewing from any chosen direction.

The *graphing* module is used to draw graphs of image motion. Displacement, velocity, and acceleration curves may be graphed for any number of individual body joints or segments. Joint motion may be presented in either linear coordinates or angular coordinates, while segment motion is presented in angular coordinates about a single segment endpoint. The data may be displayed on cartesian graphs or as full figure models.

Printed reports of the image motion data are produced by the *print* module. Data can be saved for future printing. Additionally, reports may be transferred to other systems such as spread sheets or data base programs.

The *analog* module includes a hardware interface and an analog sampling unit with program selectable gain for collection of up to 16 channels of analog input. Specialized features support the use of force plates and electromyography (EMG) measurement and analysis. Included are options for spike analysis, envelope processing, signal integration analysis, waveform analysis, and spectral analysis.

STUDIES IN PROGRESS

The Anthropometric and Biomechanics Lab (ABL) is currently involved in ongoing studies to enhance astronaut performance in a space environment. Depending on the study, all or part of the APAS may be used for data collection and analysis.

Initial investigations are in progress utilizing motion analysis, EMG, and an instrumented treadmill to measure and compare the shirtsleeved one-gravity relations between velocity, angle of inclination, skeletal muscle contraction patterns, and impact loading of the skeletal system to identical conditions in a zero-gravity environment.

An unrelated but similar investigation is in progress to determine the lower torso mobility of one of the Space Station prototype space suits in one gravity, in addition to simulated lunar and martian gravity, 1/6 and 1/3 Earth gravity respectively. The treadmill and EMG are also being incorporated into this study. The ABL is also investigating the possibility of using the APAS to determine reach envelopes of astronauts as a function of varying gravitational loads while wearing the Launch/Entry Suit (LES).

CONCLUSION

The video based motion analysis system being used by the ABL has proved to be a viable

means for collecting and analyzing human motion. A great strength of video based systems is their flexibility. The system is used in the one-gravity lab environment, in neutral buoyancy at JSC's Weightless Environment Training Facility (WETF), and in zero gravity on board NASA's KC-135. The systems are versatile and allow the operator to analyze virtually any motion that can be sequentially imaged.

REFERENCES

1. Shapiro, R. "Direct Linear Transformation Method For Three-Dimensional Cinematography." *Research Quarterly*, 49:197-205, 1978.

2. Reinsch, C. H. "Smoothing By Spline Functions." *Numerische Mathematik*, 10:177-184, 1967.

REACH PERFORMANCE WHILE WEARING THE SPACE SHUTTLE LAUNCH AND ENTRY SUIT DURING EXPOSURE TO LAUNCH ACCELERATIONS

James P. Bagian and Michael C. Greenisen
NASA Lyndon B. Johnson Space Center

Lauren E. Schafer and John D. Probe
Lockheed Engineering and Sciences Co.

Robert W. Krutz, Jr.
Krug International

INTRODUCTION

Crewmen aboard the Space Shuttle are subjected to accelerations during ascent (the powered flight phase of launch) which range up to $+3G_X$. Despite having 33 missions and nine years experience, not to mention all the time spent in development prior to the first flight, no truly quantitative reach study wearing actual crew equipment, using actual Shuttle seats and restraints has ever been done. What little information exists on reach performance while under acceleration has been derived primarily from subjective comments gathered retrospectively from Shuttle flight crews during their post mission debrief. This lack of reach performance data has resulted in uncertainty regarding emergency procedures that can realistically be performed during an actual Shuttle ascent versus what is practiced in the ground-fixed and motion-based Shuttle Simulators.

With the introduction on STS-26 of the current Shuttle escape system, the question of reach performance under launch accelerations was once again raised. The escape system's requirement that each crewman wear a Launch/ Entry Suit (LES), parachute harness, and parachute were all anticipated to contribute to a further degradation of reach performance during Shuttle ascent accelerations. In order to answer the reach performance question in a quantitative way, a photogrammetric method was chosen so that the actual reach values and associated envelopes could be captured. This would allow quantitative assessment of potential task performance impact and identify areas where changes to our Shuttle ascent emergency procedures might be required. Also, such a set of reach values would be valid for any similar acceleration profile using the same crew equipment. Potential Space Station applications of this data include predicting reach performance during Assured Crew Return Vehicle (ACRV) operations.

METHOD

Four astronaut/pilot volunteers were used as test subjects for the reach evaluations at both 1 and $3G_X$. All were veterans of one or more previous Shuttle flights and had used the crew equipment configuration under consideration numerous times before, including an actual Shuttle mission.

The LES was designed to function as a combination dry-type, anti-exposure suit and a partial pressure, high altitude protection suit. Each subject wore the LES over a set of expedition weight Capilene® underwear. A specially designed torso harness was worn over the LES and connected by quick release fasteners to a personal parachute. This parachute was worn on the crewman's back and also functioned as a seat-back cushion.

Each subject was tested during two runs on the centrifuge at Brooks Air Force Base. One run was done at $1G_X$ (lying on his back while strapped in the seat), and the other was performed at the $3G_X$ level.

The reach sweeps performed by each subject were captured by four video cameras. One

camera was secured in each corner of the centrifuge gondola and oriented for an optimum view of the subject. The four views of each recorded motion were subsequently digitized and analyzed using the Ariel Performance Analysis System, developed by Ariel Dynamics, Inc.

STATISTICAL METHODS

The data obtained from the motion analysis of left and right reach sweeps was normalized and prepared for statistical analysis. The cartesian coordinates of the left and right shoulder were noted while the subject was at rest during the $1G_X$ loading condition. These coordinates were then used as the origin for reach measurements during both the $1G_X$ and $3G_X$ sweeps. In this way, reach was normalized for each subject.

Reach was defined as the distance, in centimeters, between the shoulder and the knuckles for each coordinate. Maximum reach capability was compared in the forward, lateral and overhead (x, y, and z respectively) directions during 1 and $3G_X$ loading conditions. (Note: The measurement of lateral reach did not reflect a true maximum since all of the subjects were, at less than their full reach, able to touch the sidewalls of the gondola during the 1 and $3G_X$ exposures. Therefore, the y and Δy values were not considered for evaluation.) Changes in reach between the two G_X levels (Δx and Δz) were calculated by subtracting the $3G_X$ reach data from the $1G_X$ values. Changes in reach between the two $3G_X$ arm sweeps (Δx and Δz) were calculated by subtracting the right arm reach data from the left arm values.

Paired-t tests were used to statistically analyze reach differences in the x and z directions. This analysis was conducted on three comparisons: the left reach sweeps at 1 versus $3G_X$, the right reach sweeps at 1 versus $3G_X$, and the left versus right reach sweeps at $3G_X$. Because of the small population size (n=4), the use of the paired-t test is limited. For this reason, percent differences were also calculated for these same comparisons.

RESULTS

The results for 1 and $3G_X$ left reach sweeps are shown in Table 1. No statistically significant differences ($p < 0.05$) existed between the 1 and $3G_X$ left reach sweeps. The difference in average forward reach (Δx) for this study population was 3.3 +/- 5.0 cm. This value indicates that a greater left forward reach was achieved during the $1G_X$ loading condition. The difference in average overhead reach (Δz) was 3.9 +/- 3.4 cm. However, in this case, greater left overhead reach capability occurred during the $3G_X$ exposure.

TABLE 1.

LES Left Sweep $1G_X$ versus $3G_X$

Sub-ject	Dominant Hand	$1G_X$		$3G_X$	
		X_{dir}	Z_{dir}	X_{dir}	Z_{dir}
1	Right	28.16	55.30	25.55	55.30
2	Right	38.29	53.30	40.33	60.38
3	Right	49.66	54.18	39.55	60.42
4	Left	49.59	50.88	46.97	53.03

Sub-ject	Dominant Hand	ΔX_{dir}	ΔZ_{dir}	$\%\Delta X_{dir}$	$\%\Delta Z_{dir}$
1	Right	12.61	0	-9.27	0
2	Right	-2.04	-7.08	5.33	13.30
3	Right	10.11	-6.24	-20.36	11.52
4	Left	2.62	-2.15	-5.28	4.23

Percent differences were calculated using $1G_X$ data as the control variable (% difference = [(experimental-control)/control] x 100). While percent differences in reach for the entire population did not exceed 10%, significant (> 10%) individual differences between 1 and $3G_X$ left reach capability did exist. Specifically, subject 2 demonstrated a 13.3%

greater left overhead reach at $3G_X$ than at $1G_X$. Similarly, subject 3 displayed an 11.5% greater left overhead reach capability at $3G_X$. This same participant showed a 20.4% greater left forward reach at $1G_X$ than at $3G_X$.

No statistically significant differences were found to exist between the 1 and $3G_X$ right reach sweeps (Table 2). The Δx for the study group was 6.3 +/- 5.6 cm. That is, a greater forward reach occurred at $1G_X$ than at $3G_X$. The Δz was 6.2 +/- 7.6 cm. However, overhead reach capability was greater during the $3G_X$ loading conditions.

TABLE 2.

LES Right Sweep $1G_X$ versus $3G_X$

Sub-ject	Dominant Hand	$1G_X$		$3G_X$	
		X_{dir}	Z_{dir}	X_{dir}	Z_{dir}
1	Right	42.98	56.40	28.58	57.23
2	Right	41.74	56.21	38.22	71.62
3	Right	49.10	66.11	43.56	75.32
4	Left	43.17	73.71	41.43	72.95

Sub-ject	Dominant Hand	ΔX_{dir}	ΔZ_{dir}	$\%\Delta X_{dir}$	$\%\Delta Z_{dir}$
1	Right	14.40	-.83	-33.50	1.47
2	Right	3.52	-15.41	-8.43	27.42
3	Right	5.54	-9.21	-11.28	13.93
4	Left	1.74	.76	-4.03	-1.03

Once again, percent differences were calculated using $1G_X$ data as the control variable. There was a significant percent difference in right forward reach for the entire population. This calculation indicated that forward reach was 14.2% greater at $1G_X$ than at $3G_X$ for the entire group. Significant individual percent differences in right reach also occurred. Subject 1 demonstrated a 33.5% greater right forward reach at $1G_X$ than at $3G_X$. Subject 2 displayed a 27.4% greater right overhead reach during the 3Gx exposure. Similarly, subject 3 showed a 13.9% greater right overhead reach at the $3G_X$ level. However, this same astronaut exhibited an 11.3% greater forward reach during $1G_X$ loading conditions.

Comparison of reach at $3G_X$ in the LES revealed that a statistically significant difference ($p = .037$) did exist between left and right sweeps under $3G_X$ loading conditions (Table 3). This difference indicated that a greater right overhead reach was obtained in the LES suit. This was true for both right and left hand dominant subjects.

TABLE 3.

At $3G_X$ in LES Left versus Right Sweep

Sub-ject	Dominant Hand	LX	LZ	RX	RZ
1	Right	25.55	55.30	28.58	57.23
2	Right	40.33	60.38	38.22	71.62
3	Right	39.55	60.42	43.56	75.32
4	Left	46.97	53.03	41.43	72.95

Sub-ject	Dominant Hand	LX-RX	LZ-RZ	$\%\Delta X$	$\%\Delta Z$
1	Right	-3.03	-1.93	-10.60	-3.37
2	Right	2.11	-11.24	5.52	-15.69
3	Right	-4.01	-14.90	-9.21	-19.78
4	Left	5.54	-19.92	-11.79	37.56

Percent differences were calculated using dominant hand data as the control variable (% difference = [(nondominant)/dominant] x 100). There was a significant percent difference (17.3%) between left and right overhead reach for the entire population. This value indicates that, under $3G_X$ loading conditions, the right overhead reach was greater than the left. No other significant percent differences in mean population reach occurred. However,

significant individual differences did exist. Subject 1 showed a 10.6% greater right than left forward reach. Subject 2 demonstrated a 15.7% greater right overhead reach. Similarly, subject 3 exhibited a 19.8% greater right overhead reach. Subject 4, the only left-handed person in this group, displayed an 11.8% greater left forward reach. This participant also demonstrated a 37.6% greater right overhead reach.

SUMMARY

Since all subjects had significant previous experience using the equipment under evaluation, it is unlikely that any training effect is responsible for the results which were obtained.

The changes in reach in the +x (forward) direction were qualitatively what had been anticipated based on anecdotal reports received during Space Shuttle mission debriefings. Three of four subjects during left arm motion and four of four subjects during right arm motion experienced reduced reach capability in the +x direction at $3G_X$ versus $1G_X$. The magnitude of this change was not as great as was expected, in all cases, ranging from an improvement of 2.04 cm to a 10.11 cm decrease on the left to a 14.4 cm decrease on the right. While these differences between right and left are striking, they are not statistically significant.

It was unexpected that any reach envelopes at $3G_X$ would have been greater than that observed at $1G_X$. However, this was definitely the case in the +z (overhead) direction for three of four subjects during both left and right arm motion. The absolute range of reach difference in the +z (overhead) direction ranged from 0 to 7.08 cm on the left and -.76 to 15.41 cm on the right. These represented 13.3% and 27.4% increase in left versus right reach respectively. Operationally this would seem to indicate that any task which can be accomplished during $1G_X$ in the simulator should be achievable during actual flight.

Interestingly, there was a statistically significant difference (p = .037) between the left and right overhead reach with the right being greater. This unanticipated finding, which was unrelated to the subject's handedness, raises several points for consideration. Since the LES is symmetrically constructed, it is unlikely that it was, by itself, responsible for the asymmetry observed. The torso harness which is worn over the LES is not symmetrical (which is also the case with the parachute). It is felt that further analysis in the future of the asymmetry of the equipment may identify a course of action which will improve the left overhead reach to the point where it is equivalent to the right.

CONCLUSIONS

These data indicate that ground-based simulator training is adequate as far as verifying the feasibility of overhead activities are concerned. The same is not true of activities involving forward reach. Accordingly, to make training realistic, crewmen should be instructed that tasks involving forward reach should not be attempted during simulator runs if they exceed 66-80% of the maximum $1G_X$ forward reach capability of the crewman.

Also, more generically, this study has demonstrated the utility of using photogrammetric techniques to quantify magnitudes of reach in any direction. Further, since this data is handled and ultimately stored digitally, it is fully "portable" and can thus be used to predict reach performance in any environment where the subject is exposed to similar accelerative loads, etc.

In future work, we will merge our reach information with a graphics data base describing the Space Shuttle cockpit panels. This will allow us to find the intersection of these two data bases and represent actual panel positions reachable by a specific subject.

DEVELOPMENT OF BIOMECHANICAL MODELS FOR HUMAN FACTORS EVALUATIONS

Barbara Woolford
NASA Lyndon B. Johnson Space Center

Abhilash Pandya and James Maida
Lockheed Engineering and Sciences Company

COMPUTER MODELING OF HUMAN MOTION

Computer aided design (CAD) techniques are now well established and have become the norm in many aspects of aerospace engineering. They enable analytical studies, such as finite element analysis, to be performed to measure performance characteristics of the aircraft or spacecraft long before a physical model is built. However, because of the complexity of human performance, CAD systems for human factors are not in widespread use. The purpose of such a program would be to analyze the performance capability of a crew member given a particular environment and task. This requires the design capabilities to describe the environment's geometry and to describe the task's requirements, which may involve motion and strength. This in turn requires extensive data on human physical performance which can be generalized to many different physical configurations. PLAID is developing into such a program. Begun at Johnson Space Center in 1977, it was started to model only the geometry of the environment. The physical appearance of a human body was generated, and the tool took on a new meaning as fit, access, and reach could be checked. Specification of fields-of-view soon followed. This allowed PLAID to be used to predict what the Space Shuttle cameras or crew could see from a given point. An illustration of this use is shown in Figures 1a and 1b. Figure 1a was developed well before the mission, to show the planners where the EVA astronaut would stand while restraining a satellite manually, and what the IVA crewmember would be able to see from the window. Figure 1b is the view actually captured by the camera from the window. However, at this stage positioning of the human body was a slow, difficult process as each joint angle had to be specified in degrees.

REACH

The next step in enhancing PLAID's usefulness was to develop a way of positioning bodies by computer simulation, rather than by the engineer's inputs of joint angles. The University of Pennsylvania was contracted to perform this work. Korein (1985) developed an inverse kinematic solution for multijointed bodies. This enabled the engineer to position one "root" of the body (feet in foot restraint, or waist or hips fixed) in a specified location, and then specify what object or point in the workspace was to be touched by other parts of the body (such as place the right hand on a hand controller, and the left on a specific switch). The algorithm then attempted to find a position which would allow this configuration to be achieved. If it was impossible to achieve, due to shortness of arms or position of feet, a message would be presented giving the miss distance. This feedback enabled the engineer to draw conclusions about the suitability of the proposed body position and workspace. While this reach algorithm is extremely useful for body position, it does not enable an analyst to check an entire workspace for accessibility without specifying a large number of "reach to" points. This need has been recently met by a kinematic reach algorithm. The user specifies which joints to exercise. The algorithm then accesses an anthropometry data base giving joint angle limits, positions the proximal joint at its extreme limit, and steps the distal joint through its range of motion in a

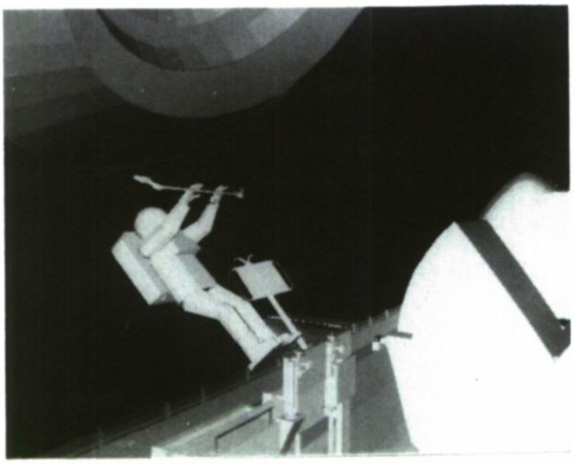


Figure 1a. PLAID rendition of crewmember restraining payload, from premission studies.

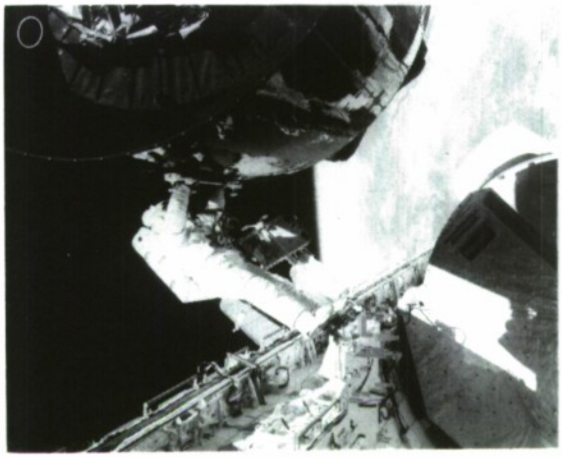


Figure 1b. Photo taken during mission from aft crew station.

number of small steps, generating a contour. The proximal joint is moved an increment, and the distal joint swung through its range of motion again. This process continues until the proximal joint reaches its other extreme limit. A three dimensional set of colored contours is thus generated which can be compared to the workstation and conclusions can be drawn. An example of this is shown in Figures 2a and 2b. In Figure 2a, a fifth percentile female is placed at the proposed foot restraint position intended to provide an eyepoint 20" from the workstation. In this position, her reach envelope falls short of the workstation. Figure 2b shows the same body and reach envelope positioned with a 16" eyepoint, in which case the woman can reach the workstation.

ANIMATION

Human performance is not static. To do useful work, the crewmembers must move their hands at least, and frequently their bodies, their tools, and their equipment. While this can be captured in a sequence of static pictures, animations are much preferred because they show all the intermediate points between the static views. Originally, PLAID animations were created by having the analyst enter every single step individually. This was highly labor intensive, and prohibitive in cost for any but the most essential conditions. However, an animation capability was created that allowed the user to input only "key frames." (A key frame is one where the velocity or direction of motion changes.) The software then smoothly interpolates 20 or 30 intermediate frame scenes, showing the continuous movement. This has many applications for both the Shuttle program and for the Space Station Freedom (SSF) program. For example, in determining where interior handholds were needed, an animation was created showing the process of moving an experiment rack from the logistics module to the laboratory module. Clearances, collisions, and points of change could be identified from the videotape. However, while the tape showed the locations for the handholds, it could not give information as to the loads the handholds would have to bear. Thus a project to model strength was begun.

BIOMECHANICS MODELING

UPPER TORSO STRENGTH

Using a Loredan, Inc. LIDO dynamometer, single joint strength data was collected for the

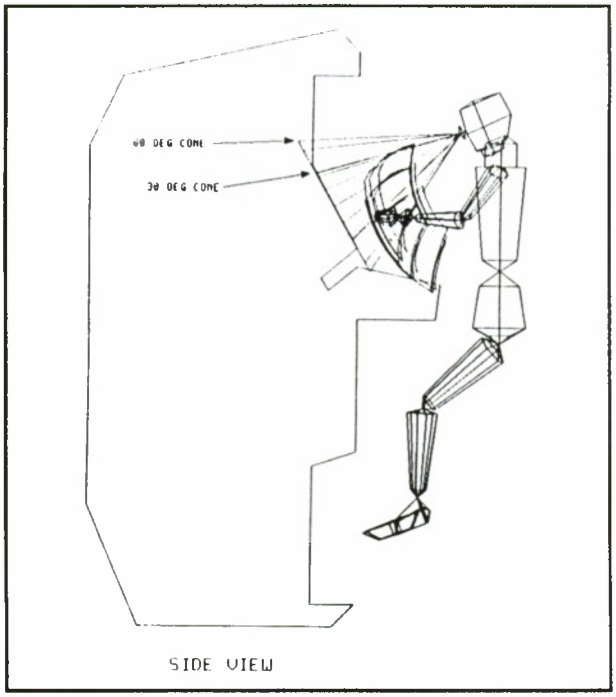


Figure 2a. Fifth percentile female positioned at workspace with 20" eyepoint. Reach contours miss the workstation.

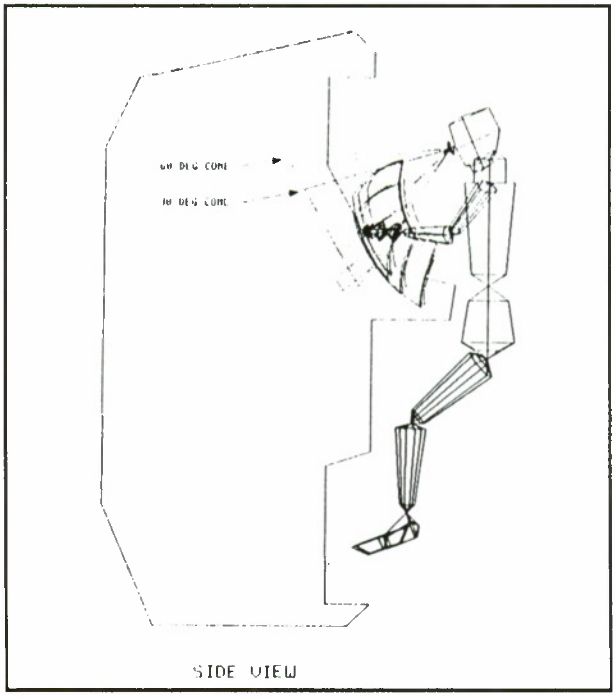


Figure 2b. Fifth percentile female positioned at workspace with 16" eyepoint. Reach contours touch workstation.

shoulder, elbow, and wrist of one individual. The data was collected in the form of (velocity, position, and strength) triplets. That is, the dynamometer was set to a selected speed, ranging from 30 deg/sec to 240 deg/sec in 30 deg/sec increments. For that speed, the subject moved his joint through its entire range of motion for the specified axis (abduction/adduction and flexion/extension). Data was collected every five degrees and a polynomial regression equation fit to the data for that velocity. The velocity was changed, and the procedure repeated. This resulted in a set of equations, giving torque in foot-pounds as a function of velocity and joint angle, for each joint rotation direction. Figure 3 shows shoulder flexion torque over a range of angles, parameterized by velocity. Figure 4 shows the data points and the equation fit for elbow flexion/extension over the range of motion at 90 deg/sec.

These regression equations were stored in tables in PLAID. To predict total strength exerted in a given position or during a given motion, the body configuration for the desired position (or sequence of positions) is calculated from the inverse kinematics algorithm. For example, the task used so far in testing is ratchet wrench push/pull. This task is assumed to keep the body fixed, and allow movement only of the arm. (As more strength data is obtained, the tasks can be made more complex.) A starting position for the wrench is established, and the position of the body is set. The angles of the arm joints needed to reach the wrench handle are then calculated. A speed of motion, indicative of the resistance of the bolt, is specified. The tables are searched, and the strength for each joint for the given velocity at the calculated angle is retrieved. The direction of the force vector is calculated from the cross products of the segments, giving a normal to the axis of rotation in the plane of rotation.

Once all these force vectors are obtained, they are summed vectorially to calculate the

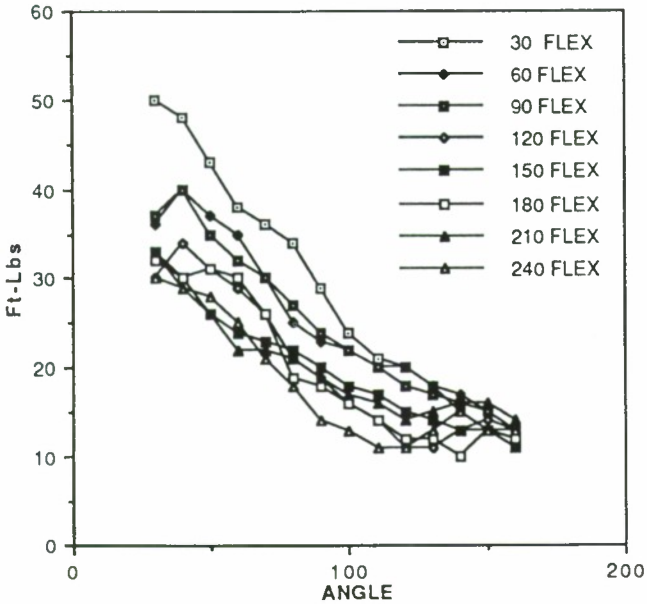


Figure 3. Shoulder flexion torque for velocities ranging from 30°/sec to 240°/sec.

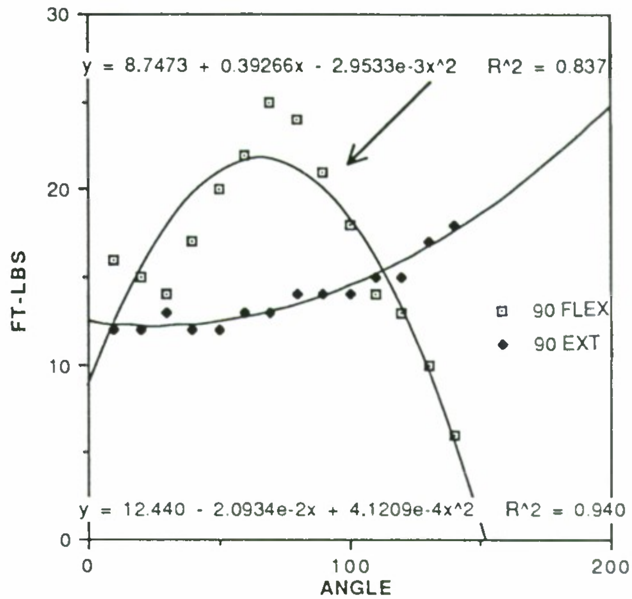


Figure 4. Raw data and regression equations for elbow flexion and extension at 90°/sec.

resultant end effector force. Currently the program displays the force for each joint and the resultant end effector force, as illustrated in Figure 5. The ratchet wrench model rotates accordingly for an angular increment. This requires a new configuration of the body, and the calculation is repeated for this new position. A continuous contour line may be generated which shows the end effector force over the entire range of motion by color coding. The model will be validated this summer. A ratchet wrench attachment for a dynamometer has been obtained, and an Ariel Motion Digitizing System will be used to measure the actual joint angles at each point in the pushing and pulling of the wrench. This will provide checks on both the validity of the positioning algorithms and of the force calculations. When this simple model is validated, more complex motions will be investigated.

The significance of this model is that it will permit strengths to be calculated from basic data (single joint rotations) rather than requiring that data be collected for each particular motion, as is done in Crew Chief (Easterly, 1989). A synthesis of the reach envelope generating algorithm and the force calculations has been achieved. The analyst can now generate reach contours which are color coded to show the amount of force available at any point within the reach envelope.

EFFECTS OF GRAVITY-LOADING ON VISION

Human vision is another important parameter being investigated in conjunction with human reach and strength. Empirical data relating maximum vision envelopes versus gravity loading have been collected on several subjects by L. Schafer and E. Saenz. This data will be tabularized in a computer readable form for use in man-modeling. Preliminary software design has begun on a vision model which will utilize this vision data to simulate a period of Space Shuttle launch where gravity loading is a major factor. This model will be able to dynamically display the vision cone of a particular individual as a function of gravity force and project that cone onto a workstation to determine if all the appropriate gauges/displays can been seen.

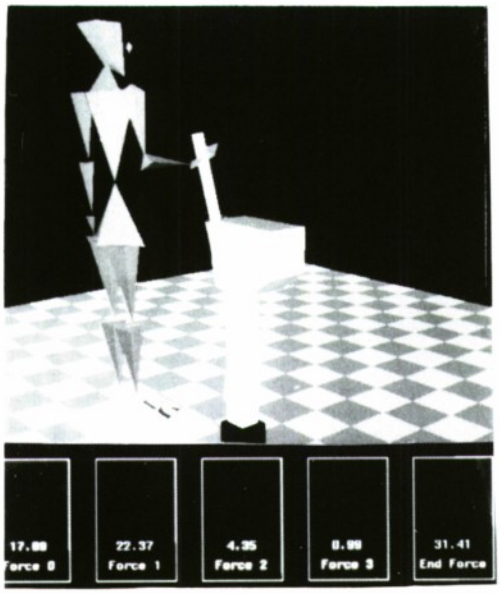


Figure 5. Body model exerting force on ratchet wrench. Joint forces and effective force at wrench are displayed as bar graphs beneath the picture.

APPLICATIONS

The biomechanical models, combined with geometric and dynamic modeling of the environment, have two major applications. The first is in equipment design. Frequently the strength or force of a crewmember is a key parameter in design specifications. For example, a manually operated trash compactor has recently been built for the Shuttle for extended duration (10-14 days) operations. This is operated by a crew member exerting force on the handle to squeeze the trash, and is seen as an exercise device as well as a trash compactor. The two key specifications needed were (1) how much force can a relatively weak crewmember exert, so the right amount of mechanical amplification can be built in, and (2) how much force could a very strong crewmember exert, so the machine could be built to withstand those forces. When the biomechanical model is completed, questions such as these can be answered during the design phase with a simulation rather than requiring extensive testing in the laboratory. In addition, the size of the equipment can be compared visually to the available storage space, and the location of foot restraints relative to the equipment can be determined. Other equipment design applications include determining the specifications for exercise equipment, determining the available strength for opening or closing a hatch or door, and determining the rate at which a given mass could be moved. The second application for a strength model is in mission planning. Particularly during extravehicular activities (EVA), crewmembers need to handle large masses such as satellites or structural elements. A complete dynamics model would enable the mission planners to view the scenes as they would be during actual operations by simulating the forces which can be exerted and the resulting accelerations of the large mass.

FUTURE PLANS

Currently the only motion modeled is a rotational motion of a wrench using only the arm, not the entire body. One step in developing a useful model is to allow the software already available for animating motion to be used to define any motion and then permit calculation of the strength available taking the entire body into account. This is a major step to accomplish, because of the many degrees of freedom in the entire human body. In order to consider the entire body in strength analysis, empirical strength data must be collected. The Anthropometry and Biomechanics Lab at the Johnson Space Center is beginning work on this project. To date, shoulder and arm strength measurements have been collected on a number of subjects. This data must be made available through the program's data base so that 5th percentile, or median, or 95th percentile strengths can be examined. This will involve another layer of data in the data base. The strength measurements for the entire body, especially torso and legs, are needed. Collecting these strength data for the individual joints at a number of angular positions and angular

velocities will be an ongoing project for some time. However, efforts have been made to automate data entry and reduction, which will result in easier data collection. Finally, the most important step is to validate the strength data. An assembly for collecting forces and angles for a ratchet wrench operation is available, and will be used to validate the compound motion of the arm. Movement of the entire body will be validated after the original data is collected, equations fit, and predictions of strength made.

REFERENCES

1. Badler, N. I., Lee, P., Phillips, C. and Otani, E. M. "The JACK Interactive Human Model." In "Concurrent Engineering of Mechanical Systems, Vol. 1." E. J. Haug, ed. *Proceedings of the First Annual Symposium on Mechanical System Design in a Concurrent Engineering Environment*, University of Iowa: Oct. 24-25, 1989.

2. Easterly, J. "CREW CHIEF: A Model of a Maintenance Technician," AIAA/NASA Symposium on the Maintainability of Aerospace Systems, July 26-27, 1989: Anaheim, CA.

3. Korein, James U. "A Geometric Investigation of Reach," MIT Press, Cambridge, MA: 1985.

ESTABLISHING A RELATIONSHIP BETWEEN MAXIMUM TORQUE PRODUCTION OF ISOLATED JOINTS TO SIMULATE EVA RATCHET PUSH-PULL MANEUVER: A CASE STUDY

Abhilash Pandya and James Maida
Lockheed Engineering and Services Company

Scott Hasson
University of Texas Medical Branch

Michael Greenisen and Barbara Woolford
NASA Lyndon B. Johnson Space Center

INTRODUCTION

As manned exploration of space continues, analytical evaluation of human strength characteristics is critical. These extraterrestrial environments will spawn issues of human performance which will impact the designs of tools, work spaces, and space vehicles.

Computer modeling is an effective method of correlating human biomechanical and anthropometric data with models of space structures and human work spaces (Figure 1). The aim of this study is to provide biomechanical data from isolated joints to be utilized in a computer modeling system for calculating torque resulting from any upper extremity motions: in this study, the ratchet wrench push-pull operation (a typical extravehicular activity task).

Established here are mathematical relationships used to calculate maximum torque production of isolated upper extremity joints. These relationships are a function of joint angle and joint velocity.

METHOD

Maximum torque data were obtained on a single subject during isolated joint movements of the shoulder, elbow, and wrist at angular velocities of 30 to 240 deg/sec at 30 deg/sec increments on the Loredan Inc. LIDO system. Data collection software tracked and stored joint angle data, as well as torque and velocity data, simultaneously. The angle versus torque data was reduced using a least squares regression algorithm to generate polynomial equations relating the two variables, torque and joint angle at various velocities.

These torque functions were then tabularized for utilization by the computer modeling system (Figure 2). The modeling system then correlated the functions with the appropriate joints in an anthropometrically correct human model. A ratchet wrench task was simulated and the force vectors generated from these isolated joint equations were then summed to yield end-effector torque.

As a preliminary step in the model validation process, isotonic (constant load) maximum torque data were collected for the ratchet wrench push-pull operation. Plans to collect more controlled (restricted motions) isokinetic (constant velocity) ratchet wrench data to match model outputs are in progress.

RESULTS

Second order regression equations relating joint angle to end-effector torque of the shoulder, elbow and wrist in all axes, and directions at various velocities were established. The data indicated a relationship between the allowed velocity (i.e., decreased velocity was proportional to increased resistance) and the torque generated. As indicated in Figure 3, the maximum torque generated decreases as the velocity increases.

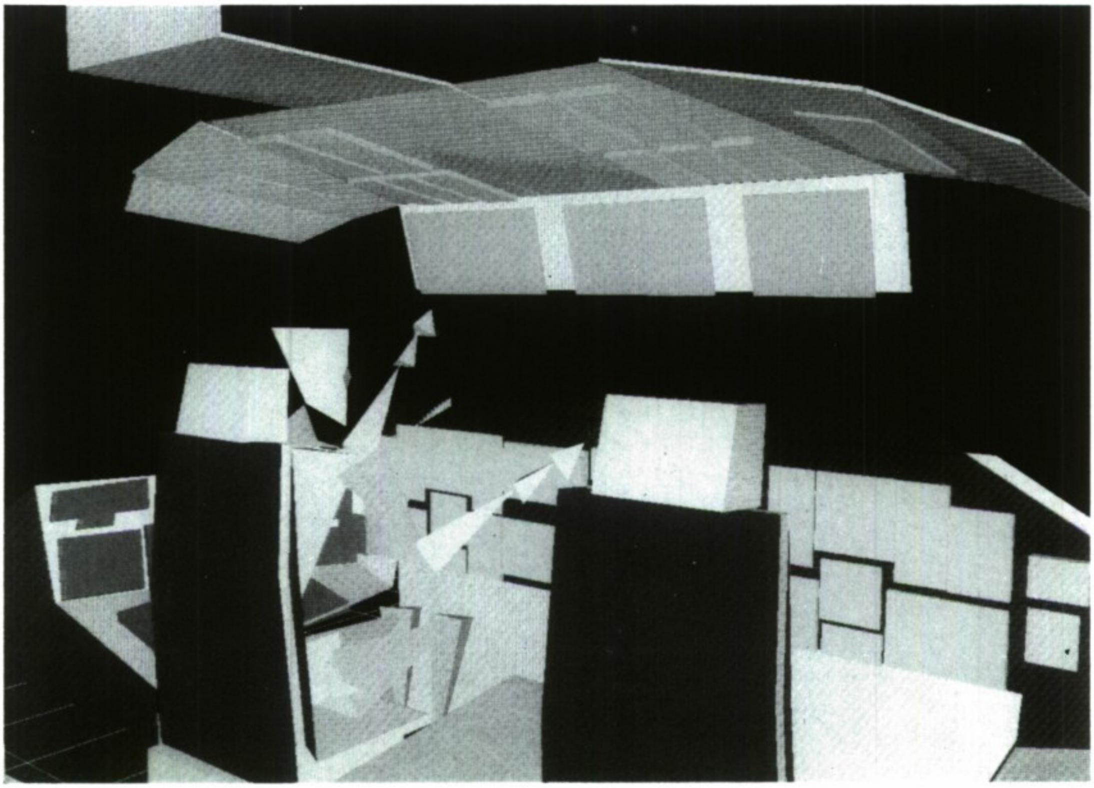


Figure 1. Man model representation in a work space.

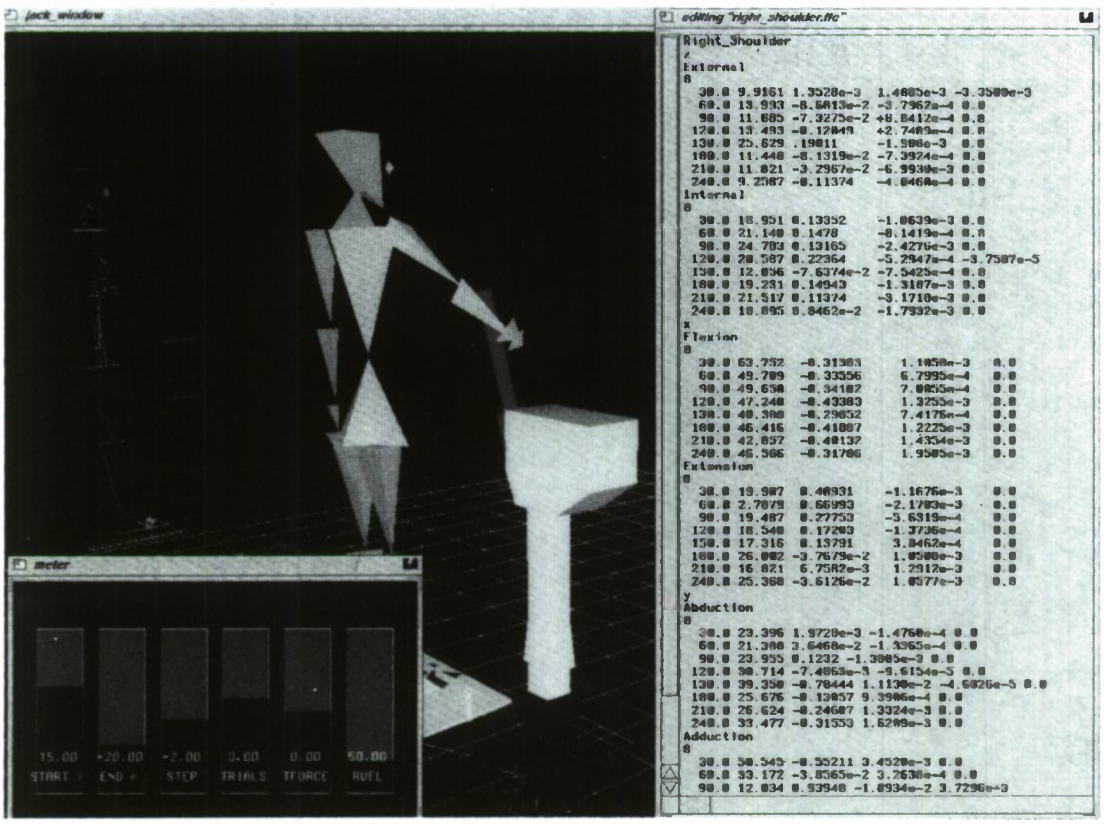


Figure 2. Torque function regression coefficients incorporated into man model.

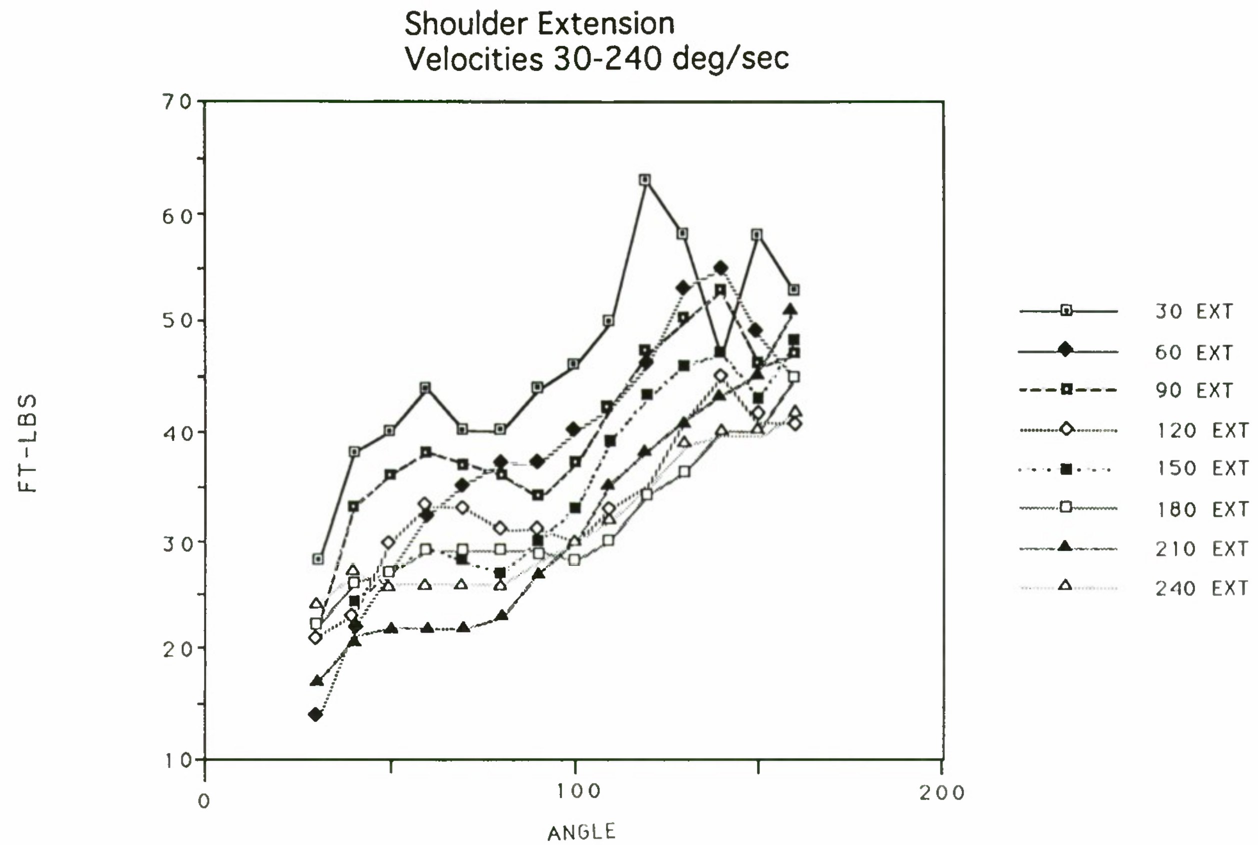


Figure 3. Shoulder regression equation at all measured velocities.

All the isolated joint relationships were coded into a flexible and interactive computer graphics model. This model allowed alteration of the initial position and joint angles of the human figure relative to the ratchet wrench. This flexibility allowed one to gauge the effects of body orientation on torque generated.

The calculation for torque generated was for the isokinetic ratchet wrench motion. Model validation data for this configuration is now being collected.

CONCLUSION

It has been demonstrated that a computer model may be a viable method to calculate torque resulting from arbitrarily complex motions. Using regression equations derived from empirically measured torques for isolated joints, end-effector torque was calculated and displayed for an isokinetic ratchet wrench procedure (Figure 4).

For initial validation efforts, isotonic data on the ratchet wrench were collected. Because of the uncontrolled ratchet velocities in the isotonic measurements, model calculations (based on isokinetic configuration) were not acceptably accurate (up to 40% lower). An accurate validation and refinement of the model is contingent upon collection of very controlled (restricted motion) isokinetic data (constant velocity) of the ratchet wrench motion for more subjects.

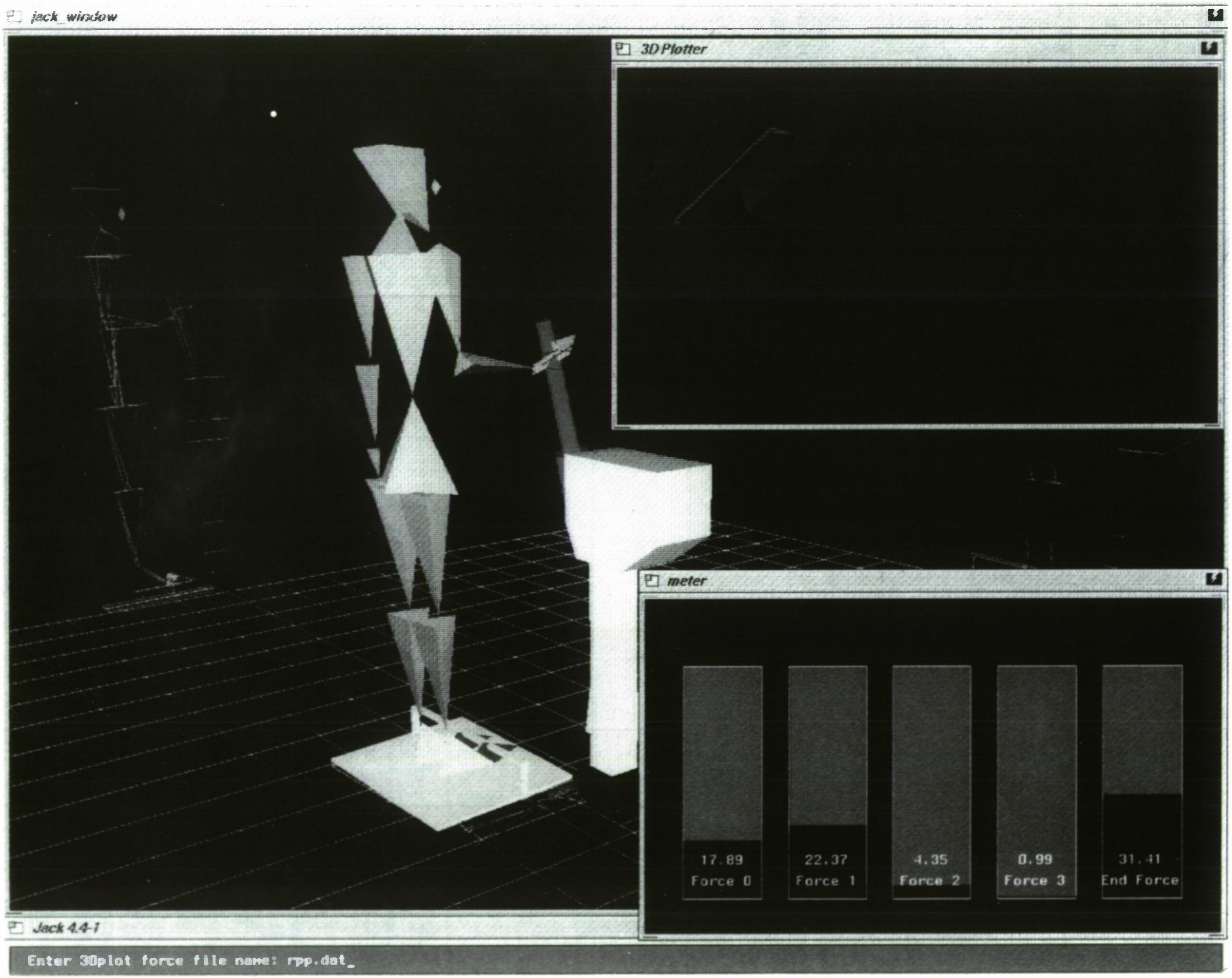


Figure 4. Display of joint and end-effector torques for the ratchet wrench task push-pull task.